

目录

| | |
|-------------------------------------|----|
| 一、问题的提出..... | 1 |
| 二、问题的分析..... | 2 |
| 三、符号约定..... | 3 |
| 四、模型的假设..... | 4 |
| 五、问题 I—二体轨道力学和运动方程的模型..... | 4 |
| 5.1 二体问题和运动方程：..... | 4 |
| 5.2 进一步简化的运动方程：..... | 5 |
| 5.3 轨道的几何方程的进一步探究：..... | 5 |
| 六、问题 II-1—解析解模型..... | 7 |
| 6.1 解析解模型的分析：..... | 7 |
| 6.2 解析解模型的求解：..... | 8 |
| 七、问题 II-1—求最小组数的最小二乘法模型..... | 9 |
| 7.1、离散格式下的求最小 m 值的算法..... | 9 |
| 7.2、利用解析解求最小 m 值的算法..... | 12 |
| 7.3 模型的求解与检验..... | 13 |
| 八、问题 II-3—高精度差商的最小二乘法模型..... | 16 |
| 8.1 解析解模型的求解和讨论..... | 16 |
| 8.2 考虑了高精度差商的最小二乘法模型的分析 and 求解..... | 18 |
| 九、问题 II-3—变分伴随同化模型..... | 27 |
| 9.1 问题的分析..... | 27 |
| 9.2 问题的求解：..... | 32 |
| 9.3 对模型误差敏感度的检验：..... | 37 |
| 十、问题 II-4—中央差商的最小二乘法模型..... | 39 |
| 10.1 分析和模型修改..... | 39 |
| 10.2 结果讨论..... | 40 |
| 10.3 对模型误差敏感度的检验：..... | 42 |
| 十一、问题 II-4—变分伴随同化模型..... | 45 |
| 11.1 问题的分析..... | 45 |
| 11.2 问题的求解：..... | 45 |
| 11.3 对模型误差敏感度的检验：..... | 48 |
| 十二、模型的推广..... | 49 |
| 12.1 背景知识：..... | 49 |
| 12.2 考虑捕获量这个外界因素的影响：..... | 50 |
| 12.3 考虑更为广泛的捕食系统：..... | 51 |
| 十三、模型的优缺点分析..... | 54 |
| 十四、关于本问题的几点讨论..... | 54 |
| 十五、模型的扩展..... | 55 |

一、问题的提出

包括“神舟六号”载人航天宇宙飞船、人造地球卫星等航天器围绕地球在轨运行的过程中，要受到很多力的作用，其中主要的是地球万有引力和航天器发动机作用力。

一：考虑航天器在仅受到地球万有引力、航天器自身发动机作用力的作用下作平面运动，将地球和航天器视为质点，试建立航天器运动的数学模型（只要列出模型，不要求解）。

显然这样的数学模型在精度上是远远不能满足实际需要的，在其他要求精确制导等有关高科技的实际问题中，我们都面临着类似的问题：我们必须建立高精度的数学模型，必须高精度地估计模型中的大批参数，因为只有这样的数学模型才能解决实际问题，而不会出现差之毫厘，结果却失之千里的情况。这时所建立数学模型的精度就成了数学模型的生命线。例如上述问题中的航天器还要受到地球质量分布不均匀所引起的摄动力，大气阻力，日、月及其它星球的摄动引力的影响，以及航天器发动机为调整航天器自身姿态运作时作用力的影响。这样不但数学模型十分复杂，而且在这些数学模型中还要涉及到许多重要的参数，如地球的引力场模型就有许多待定参数。不仅如此，在对航天器进行测量时，还涉及到观测站的地理位置以及设备的系统误差等参数。为此人们要设法利用长期积累的丰富的观测资料，高精度确定这些重要的参数。

由于航天器的问题太复杂，下面本题仅考虑较简单的确定高精度参数问题。

假设有一个生态系统，其中含有两种生物，即：A 生物和 B 生物，其中 A 生物是捕食者，B 生物是被捕食者。假设 t 时刻捕食者 A 的数目为 $x(t)$ ，被捕食者 B 数目为 $y(t)$ ，它们之间满足以下变化规律：

$$\begin{cases} x'(t) = x(t)[\alpha_1 + \alpha_2 y(t)] \\ y'(t) = y(t)[\alpha_3 + \alpha_4 x(t)] \end{cases}$$

初始条件为：

$$\begin{cases} x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$$

其中 $\alpha_k (1 \leq k \leq 6)$ 为模型的待定参数。

通过对此生态系统的观测，可以得到相关的观测数据。观测数据的格式依次为：

观测时刻 t_j 、A 生物数目 $x(t_j)$ 、B 生物数目 $y(t_j)$

二：请利用有关数据，解决以下问题：

1) 在观测数据无误差的情况下，若已知 $\alpha_2 = \frac{1}{5}$ ，求其它 5 个参数

$\alpha_k (k=1,3,4,5,6)$ ？有关数据见数据文件：DATA1.TXT

- 2) 在观测数据无误差的情况下, 若 α_2 也未知, 问至少需要多少组观测数据, 才能确定参数 $\alpha_k (1 \leq k \leq 6)$? 有关数据见数据文件: DATA1.TXT
- 3) 在观测资料有误差 (时间变量不含有误差) 的情况下, 请分别利用观测数据 DATA2.TXT 和 DATA3.TXT, 确定参数 $\alpha_k (1 \leq k \leq 6)$ 在某种意义下的最优解, 并与仿真结果比较, 进而改进你们的数学模型。
- 4) 假设连观测资料的时间变量也含有误差, 试利用数据 DATA4.TXT, 建立数学模型, 确定参数 $\alpha_k (1 \leq k \leq 6)$ 在某种意义下的最优解。

二、问题的分析

问题 I 中建立二体轨道力学和运动方程的模型来解决此问题。分别分析了二体问题和运动方程, 进一步简化的运动方程, 以及讨论了轨道的几何方程和轨道可能的形状。

问题 II 实际包含了四个小问: (1) 由于 α_2 已知, 我们主要是用求解解析式的方法来直接求解出 $\alpha_k (k = 1, 3, 4, 5, 6)$ (2) 采用最小二乘法来求解。假设有 m 组数据 $(x_i, y_i), i = 1, \dots, m$, 分解析解和离散差分格式这两种情况来讨论。离散情况下: 首先将原微分方程组按某种差分格式离散化, 然后将 m 组观测数据离散后的差分格式, 构造以参数为未知变量的二元一次线性方程组, 根据最小二乘法的原理, 结合矩阵的相关理论, 讨论解存在、唯一、稳定的条件, 从而确定求解参数需要观测资料的组数 m 。解析解的利用: 从原始方程组求出包含未知参数的解析解, 同样将 m 组数据代入解析解, 构造出超定的二元一次方程组, 然后用同 (1) 的方法讨论, 最终确定最小的 m 。(3) 先用的是高精度差商的最小二乘模型, 求出最优解, 用仿真解与观测资料进行比较, 效果非常好。由于此问要考虑观测资料的误差, 所以就必须对模型的误差敏感度进行讨论, 我们发现高精度差商的最小二乘模型的敏感度也是比较好的。然后我们又用了变分伴随同化模型, 在精确度和敏感度上面做了进一步的改进。(4) 与 (3) 所用方法基本一致, 也是先用中央差商的最小二乘模型求最优解, 再用变分伴随同化模型作进一步的改进。虽然此时不仅要考虑观测资料的误差还要考虑时间误差, 但是模型的稳定性还是很好。

三、符号约定

| | |
|---------------------------------|-----------------------|
| $x(t)$ | ----- t 时刻捕食者 A 的数目 |
| $y(t)$ | -----被捕食者 B 数目 |
| $\alpha_i (i = 1, 2, \dots, 6)$ | -----种群模型中的待定参数 |
| t_j | -----观测时间序列 |
| m | -----观测资料的数据组数 |
| F_g | -----地球和航天器的万有引力 |
| G | -----万有引力常数 |
| M | -----地球质量 |
| m | -----航天器的质量 |
| r | -----地球与航天器之间的距离 |
| K | -----积分常数 |
| A | -----系数矩阵 |
| b | -----线性方程组右端的向量 |
| J | -----目标函数 |
| k_i, l_i | -----中间迭代值 |
| h | -----等距的时间步长 |
| $randn$ | -----0 到 1 之间的随机数 |
| N | -----用来表示数据精度的正数 |
| α^0 | -----参数的初值 |
| d^k | -----搜索方向 |
| ρ^k | -----搜索步长 |
| w | -----权重系数 |
| (P, Q) | -----伴随模式中的伴随变量 |
| x^{obs}, y^{obs} | -----观测资料数据 |
| x_{mean}, y_{mean} | ----- A, B 种群的平均值 |

四、模型的假设

- 1 考虑航天器在仅受到地球万有引力、航天器自身发动机作用力的作用下作平面运动；
- 2 将地球和航天器视为质点；
- 3 假设题目给定的数据资料完全正确合理，具有实际意义；

五、问题 I —二体轨道力学和运动方程的模型

主要用此模型来解决题目中问题 I 。

5.1 二体问题和运动方程：

无论是航天器还是天然天体，它们在运动中的任何给定时刻，均会受到多个周围天体的万有引力作用，甚至还有其他的力，例如阻力、推力和太阳辐射压力等的作用。本文中已经进行了简化，只考虑地球和航天器两个质体，只考虑地球万有引力和航天器发动机作用力。

分析的第一步应选择一个适于物体运动的坐标系。这件事做起来并不容易，因为所选的任何坐标系其惯性特征都存在某种程度的不确性。为不失一般性，假定存在某个合适的惯性坐标系 $O'XYZ'$ ，分析所受得力：

$$F = F_g + F_2 \quad (\text{其中 } F_g = -\frac{GMm}{r^3} \vec{r})$$

$$\text{所以: } \frac{d}{dt}(mv) = F$$

把对时间的导数展开，得到：

$$m \frac{dv}{dt} + v \frac{dm}{dt} = F \quad (1-1)$$

航天器可能不断排出某些质量以产生推力，所以 $v \frac{dm}{dt}$ 不为零，各项除以 m ，可得航天器的一般运动方程为：

$$\begin{cases} \ddot{r} = \frac{F}{m} - r \frac{\dot{m}}{m} \\ r|_{t=0} = r_0, \dot{r}|_{t=0} = \dot{r}_0, \end{cases} \quad (1-2)$$

\dot{r} 和 \ddot{r} 分别为航天器相对于惯性坐标系 $O'XYZ'$ 的速度矢量和加速度矢量。

m 和 \dot{m} 分别是物体的质量和质量随时间变化。

5.2 进一步简化的运动方程:

首先, 作两个简化假设:

- (1) 物体为球对称的, 这样就可以把物体看作质量集中在其中心。
- (2) 除了沿两物体中心连线作用的引力外, 没有其他外力和内力作用。

由 5.1 可知此时方程简化为:

$$\ddot{\vec{r}} = -\frac{G(M+m)}{r^3} \vec{r} \quad (1-3)$$

又因为航天器的质量 m 将比天体的质量 M 小得多, 因此有: $G(M+m) \approx GM$

所以运动方程又转化成:

$$\ddot{\vec{r}} + \frac{\mu}{r^3} \vec{r} = 0 \quad (\text{其中 } \mu = GM) \quad (1-4)$$

5.3 轨道的几何方程的进一步探究:

根据 5.2 得到的二体的简化方程 (1-4), 该方程在形式上十分简单, 但已经完全确定了相应轨道的形状和大小。

将方程式(1-4)两边同时与 \vec{h} 叉乘, 有

$$\ddot{\vec{r}} \times \vec{h} = -\frac{\mu}{r^3} \vec{r} \times \vec{h} = \frac{\mu}{r^3} \vec{h} \times \vec{r} \quad (1-5)$$

考虑到 h 守恒和矢量运算规则:

$(\vec{a} \times \vec{b}) \times \vec{c} = \vec{b}(\vec{a} \cdot \vec{c}) - \vec{a}(\vec{b} \cdot \vec{c})$ 及 $\vec{r} \cdot \dot{\vec{r}} = r \dot{r}$, 所以

$$\frac{d}{dt}(\dot{\vec{r}} \times \vec{h}) = \ddot{\vec{r}} \times \vec{h} + \dot{\vec{r}} \times \dot{\vec{h}} = \ddot{\vec{r}} \times \vec{h}$$

$$\frac{\mu}{r^3} \vec{h} \times \vec{r} = \frac{\mu}{r^3} (\vec{r} \times \vec{v}) \times \vec{r} = \frac{\mu}{r} \vec{v} - \frac{\mu \dot{r}}{r^2} \vec{r} = \mu \frac{d}{dt} \left(\frac{\vec{r}}{r} \right)$$

于是，可以将式(5)改写为

$$\frac{d}{dt} (\vec{r} \times \vec{h}) = \mu \frac{d}{dt} \left(\frac{\vec{r}}{r} \right)$$

两边积分得

$$\vec{r} \times \vec{h} = \mu \frac{\vec{r}}{r} + \vec{B}$$

这里 \vec{B} 是积分常矢量。用 \vec{r} 点乘该式就得到标量方程

$$\vec{r} \bullet \dot{\vec{r}} \times \vec{h} = \vec{r} \bullet \mu \frac{\vec{r}}{r} + \vec{r} \bullet \vec{B}$$

因为 $\vec{a} \bullet \vec{b} \times \vec{c} = \vec{a} \times \vec{b} \bullet \vec{c}$, $\vec{a} \bullet \vec{a} = a^2$ 总成立，所以得到

$$h^2 = \mu r + r B \cos \nu$$

式中， ν 为常矢量 \vec{B} 和矢径 \vec{r} 之间的夹角。解出 r 得轨道的几何方程为

$$r = \frac{h^2 / \mu}{1 + (B / \mu) \cos \nu} \quad (1-6)$$

令 $p = h^2 / \mu$, $e = B / \mu$, 则上式即为

$$r = \frac{p}{1 + e \cos \nu} \quad (1-7)$$

显然，轨道的几何方程是一个圆锥曲线的极坐标方程，中心引力体质心即为极坐标的原点，位于一焦点上，极角 ν 为 \vec{r} 与圆锥曲线上离焦点最近的一点与焦点连线间的夹角，常数 p 称为“半正焦弦”，常数 e 称为“偏心率”，它确

定了方程式(7)表示的圆锥曲线的类型。

至此，可以把航天器的轨道运动总结如下：

(1)圆锥曲线族(圆、椭圆、抛物线、双曲线)为二体问题中的航天器惟一可能的运动轨道。

(2)中心引力体中心必定为圆锥曲线轨道的一个焦点。

(3)当航天器沿着圆锥曲线轨道运动时，其比机械能(单位质量的动能和势能之和)保持不变。然而，动能和势能这两种形式的能量却可以相互转换。这意味着，当航天器高度增高(r 增加)时，其速度必定变慢，当 r 减少时，速度加快。正是以这种方式，使 ε 保持不变。

(4)航天器绕中心引力体运动，当 \bar{r} 和 \bar{v} 沿轨道变化时，比角动量 \bar{h} 保持不变。

(5)轨道运动总是处在一个固定于惯性空间的平面内。

六、问题 II -1—解析解模型

问题要求当观测数据无误差时，并且已知 α_2 ，求其他五个参数。

6.1 解析解模型的分析：

由题目所给的微分方程组模型：

$$\begin{cases} \frac{dx}{dt} = x(t)[\alpha_1 + \alpha_2 y(t)] \\ \frac{dy}{dt} = y(t)[\alpha_3 + \alpha_4 x(t)] \end{cases}$$

将以上方程组的自变量 t 消去，得到的 y 与 x 之间的函数关系为：

$$\frac{dy}{dx} = \frac{\alpha_3 y + \alpha_4 xy}{\alpha_1 x + \alpha_2 xy}$$

采用分离变量方法得到其通解为：

$$\alpha_1 \ln y + \alpha_2 y - \alpha_3 \ln x - \alpha_4 x = K \quad (2-1)$$

其中 K 为任意常数，如果给定了初始条件 (x_0, y_0) ，则 K 为定值。即：

$$K = \alpha_1 \ln y_0 + \alpha_2 y_0 - \alpha_3 \ln x_0 - \alpha_4 x_0 \quad (2-2)$$

由上面两式可得：

$$\alpha_1 \ln\left(\frac{y}{y_0}\right) + \alpha_2 (y - y_0) - \alpha_3 \ln\left(\frac{x}{x_0}\right) - \alpha_4 (x - x_0) = 0 \quad (2-3)$$

其中已知 $\alpha_2 = 0.2$ ，(2-3)式可以化为：

$$\alpha_1 \ln\left(\frac{y}{y_0}\right) - \alpha_3 \ln\left(\frac{x}{x_0}\right) - \alpha_4 (x - x_0) = -0.2(y - y_0)$$

任意时刻的观测数据 (x_i, y_i) 都应该满足上式。假设有 $m+1$ 组观测数据(包括初始时刻)，利用上式就可以得到如下形式的方程组：

$$A_{m \times 3} \alpha_{3 \times 1} = B_{m \times 1}$$

其中 $\alpha_{3 \times 1} = (\alpha_1, \alpha_3, \alpha_4)^T$ 为待求参数， A, B 为系数矩阵。通过求解上述方程组就可以得到要求的参数

6.2 解析解模型的求解：

利用 data1.txt 中提供的 6 组观测数据，可以得到式中的系数矩阵 $A_{5 \times 3}, B_{5 \times 1}$ ，求解得到：

$$\alpha_1 = -2, \alpha_2 = 0.2, \alpha_3 = 12, \alpha_4 = -1$$

$$\text{又因为} \begin{cases} \alpha_5 = x(t_0) \\ \alpha_6 = y(t_0) \end{cases} \text{ 所以: } \alpha_5 = 10, \alpha_6 = 60。$$

这样利用初值，由(2-2)式可得到： $K = 13.82$ ，这样就可以求出关于 x 和 y 的关系式：

$$-2 \ln y + 0.2y - 12 \ln x + x + 13.82 = 0 \quad (2-4)$$

我们就可以求出 $t = 0.1$ 到 $t = 0.5$ 时间仿真的 (\bar{X}, \bar{Y}) ，跟真实的 (X, Y) 进行比较，如下图所示：

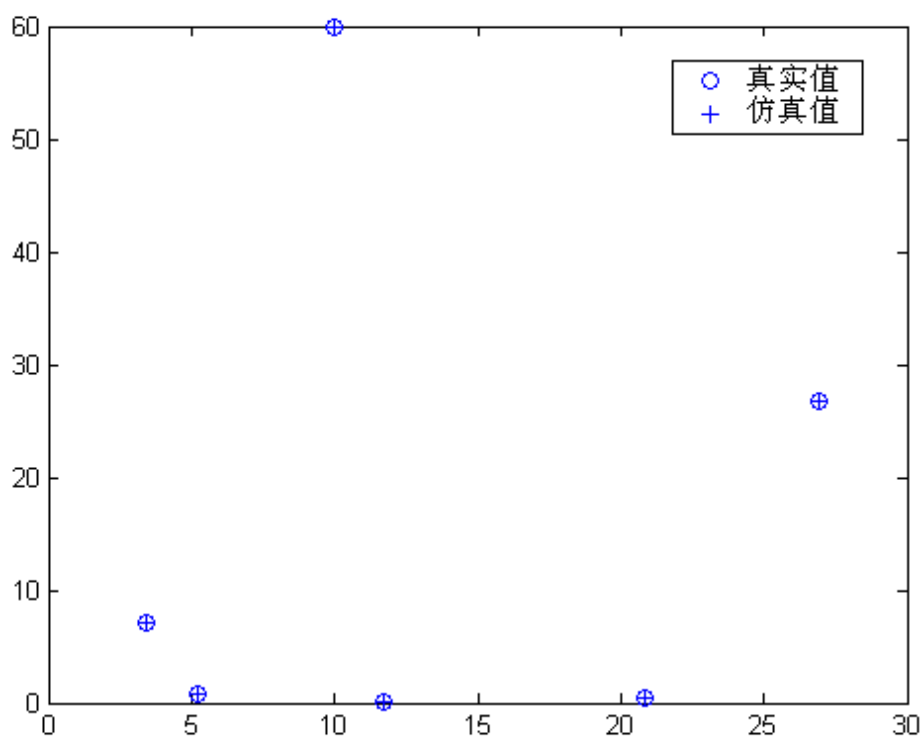


图 1 解析解模型仿真结果

从图上看，可以观测值和仿真值吻合很好，所以我们这个 $\alpha_k (k = 1, 2, 3, 4, 5, 6)$ 参数求得还是很好的。

七、问题 II -1—求最小组数的最小二乘法模型

分两种情形进行讨论。

7.1、离散格式下的求最小 m 值的算法

原微分方程组为：

$$\begin{cases} x'(t) = x(t)[\alpha_1 + \alpha_2 y(t)] \\ y'(t) = y(t)[\alpha_3 + \alpha_4 x(t)] \end{cases}$$

$$\begin{cases} x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}$$

将方程组左边的导数项离散化，采用向后差分格式：

$$\begin{cases} x(t_i)\alpha_1 + x(t_i)y(t_i)\alpha_2 = \frac{x(t_i) - x(t_{i-1})}{t_i - t_{i-1}} \\ y(t_i)\alpha_3 + x(t_i)y(t_i)\alpha_4 = \frac{y(t_i) - y(t_{i-1})}{t_i - t_{i-1}}. \quad (i = 2 \dots m) \\ \alpha_5 = x(0) = x(t_1) \\ \alpha_6 = y(0) = y(t_1) \end{cases}$$

将 α_i 看作未知变量，可以写出此方程组的矩阵形式： $A\bar{\alpha} = b$ ，其中

$$A = \begin{bmatrix} x(t_2) & x(t_2)y(t_2) & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x(t_m) & x(t_m)y(t_m) & 0 & 0 & 0 & 0 \\ 0 & 0 & y(t_2) & x(t_2)y(t_2) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & y(t_m) & x(t_m)y(t_m) & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$A_{(2n+2) \times 6}$ 由分块矩阵组成，左上的两个分块阵都有 n 行数据（ $n = m - 1$ ，即将第一个数据作为初始值）， b 同样可以写成如下格式：

$$b = \begin{bmatrix} \frac{x(t_2) - x(t_1)}{t_2 - t_1} \\ \vdots \\ \frac{x(t_m) - x(t_1)}{t_m - t_1} \\ \frac{y(t_2) - y(t_1)}{t_2 - t_1} \\ \vdots \\ \frac{y(t_m) - y(t_1)}{t_m - t_1} \\ x(t_1) \\ y(t_1) \end{bmatrix}_{(2n+2) \times 1}$$

而，

$$\alpha = [\alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5 \quad \alpha_6]^T$$

因此问题转化为要求 $A\bar{\alpha} = b$ 存在唯一稳定的非零解时至少需要多少组数据，即 m 最小为多少时可以求出 α 。一般情况下， $A\bar{\alpha} = b$ 所表示的是一个超定的线性方程组（方程个数大于未知数的个数）。在最小二乘的意义下，可以求出

该线性方程组的一组最优解，设为 $\bar{\alpha}_p$ ，使其满足误差

$$J = \|A\bar{\alpha}_p - b\|^2 = \min!$$

Step1:证明 “ $J = \|A\bar{\alpha}_p - b\|^2 = \min!$ ” 等价于 “ $A^T A\bar{\alpha}_p = A^T b$ ”

证明：记 $\langle \rangle$ 为欧氏空间的内积，构造泛函：

$$J = \frac{1}{2} \langle A\bar{\alpha}_p - b, A\bar{\alpha}_p - b \rangle$$

可以推出

$$\begin{aligned} \delta J &= \langle A\delta\bar{\alpha}_p, A\bar{\alpha}_p - b \rangle \\ &= \langle \delta\bar{\alpha}_p, A^T A\bar{\alpha}_p - A^T b \rangle = 0 \end{aligned}$$

利用 $\delta\bar{\alpha}_p$ 的任意性，得出

$$A^T A\bar{\alpha}_p = A^T b, \text{ 得证。}$$

Step2:证明 $A^T A\bar{\alpha}_p = A^T b$ 可解

证明：设 $A^T A\bar{\alpha}_p = A^T b$ 的相应齐次方程为 $A^T A\bar{y} = 0$ ，只需证明 $\langle A^T b, \bar{y} \rangle = 0$ ，即

$A^T b$ 垂直于 \bar{y} 。又因为

$$\langle A^T b, \bar{y} \rangle = \langle A^T A\bar{\alpha}_p, \bar{y} \rangle = (\bar{\alpha}_p, A^T A\bar{y}) = 0$$

其中， $()$ 为 R^6 中内积，于是 $A^T A\bar{\alpha}_p = A^T b$ ，得证。

Step3: 证明 $A^T A\bar{\alpha}_p = A^T b$ 的解不唯一

证明：如果 $J = \|A\bar{\alpha}_p - b\|^2 = \min!$ 存在一个解 x^* ，而 $A\bar{\alpha} = 0$ 有非零解 \hat{x} ，则有：

$J = \|A(x^* + t\hat{x}) - b\|^2 = \min!$ 仍然是成立的。不妨记该问题的最小二乘解集为

χ ，从该解集中取一个代表元 x^+ 作为问题的解， $x^+ = \min_{x \in \chi} \|x\| = A^+ b$ （广义逆）。

由 Step1 的证明，方程组有非零解的条件是 $|A^T A| \neq 0$ ，此时解为：

$$\bar{\alpha}_p = (A^T A)^{-1} A^T b = A^{-1} b。$$

若对 m 从 2 开始往上搜索，求出每种情况下的 $|A^T A|$ 和 α ，并计算出每个 m 值对应的模拟误差，由此综合分析比较，使得行列式 $|A^T A| \neq 0$ 同时模拟误差在容

许的范围之内的 m 即为所求。

7.2、利用解析解求最小 m 值的算法

首先来求原种群模型微分方程组的解析解，将原微分方程组改写为；

$$\begin{cases} \frac{dx}{dt} = x(\alpha_1 + \alpha_2 y) \\ \frac{dy}{dt} = y(\alpha_3 + \alpha_4 x) \end{cases}$$

两式相比，化简为

$$\frac{\alpha_3 + \alpha_4 x}{x} dx = \frac{\alpha_1 + \alpha_2 y}{y} dy$$

两边同时积分，移项得

$$\alpha_1 \ln y + \alpha_2 y - \alpha_3 \ln x - \alpha_4 x = K$$

结合初始条件 $\alpha_5 = x(1), \alpha_6 = y(1)$

其中， K 为积分常数,与初始值有关

$$K = \alpha_1 \ln y(1) + \alpha_2 y(1) - \alpha_3 \ln x(1) - \alpha_4 x(1) = \text{const}$$

将 m 组数据分别代入该解析解，构造出包含未知参数的二元一次线性方程组，其矢量形式记为： $A_1 \bar{\alpha} = b_1$ ，此时系数矩阵和等式右边的向量变为如下：

$$A = \begin{bmatrix} \ln y(t_2) & y(t_2) & -\ln x(t_2) & x(t_2) & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \ln y(t_m) & y(t_m) & -\ln x(t_m) & x(t_m) & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$b_1 = \begin{bmatrix} K \\ K \\ \vdots \\ x(t_1) \\ y(t_2) \end{bmatrix}$$

方程组有非零解的条件是 $|A_1^T A_1| \neq 0$ ，此时 K 可以赋值为 1，它并不影响行列式的求解。此时解为： $\bar{\alpha}_p = (A_1^T A_1)^{-1} A_1^T b = A_1^{-1} b_1$ 。同上，对 m 从 2 开始往

上搜索，求出每种情况下的 $|A^T A|$ 和 α ，并计算出每个 m 值对应的模拟误差，由此综合分析比较，使得行列式 $|A^T A| \neq 0$ 同时模拟误差在容许的范围之内的 m 即为所求。

7.3 模型的求解与检验

根据上面的算法，利用 data1.txt 的对上述两种情况进行检验。

7.3.1、离散情况下求最小 m 值

表 2-1 是 m 分别取 2, 3, ... 等时对应的系数行列式 $|A^T A|$ ，二元一次线性方程组的最小二乘解 α_i 和对应误差 J 的情况。

由表分析可知， $m=2$ 时， $|A^T A|$ 的数量级是 10 的-31 次方，近似的可以认为行列式的值为零，因此 $m=2$ 对应的线性方程组无解。当 $m \geq 3$ 后，行列式均是大于零的，即此种情况下线性方程组的是存在的，但不是每个存在的解都是稳定的，为找到综合意义上的最优解，还要比较通过各解模拟的模拟误差，模拟误差越小，说明其对应的解越稳定，此时的 m 就是要找的最小组数。

而表 2-1 中 $m=3$ 与 $m=4$ 、5 比较知（ $m=6$ 时由于差分格式的误差和 data1.txt 数据的原因波动较大，可暂不比较）， $m=3$ 时误差最小，说明提供三组数据时线性方程组即存在唯一，而且能在一定程度上体现出最小二乘法的稳定性。所以，在离散的情况下至少需要 3 组数据就可以确定参数 α 。

表 1 离散情况下，不同 m 值时的 $|A^T A|$ 、 α 和误差

| 参 数 组 值 | $ A^T A $ | α_1 | α_2 | α_3 | α_4 | α_5 | α_6 | 模拟误差 |
|------------------|--------------|------------|------------|------------|------------|------------|------------|--------------|
| $m = 6$ | 6.122562e+15 | 0.911920 | 0.025215 | 1.6791 | 0.009381 | 10 | 60 | 3.752589e+2 |
| $m = 5$ | 9.520669e+8 | 1.347800 | -0.90681 | 16.239 | -4.213100 | 10 | 60 | 1.824217e+10 |
| $m = 4$ | 8.733238e+8 | 1.621800 | -0.82853 | 16.545 | -4.320400 | 10 | 60 | 4.308100e+5 |
| $m = 3$ | 5.188582e+6 | 0.400190 | -0.74352 | 359.37 | -104.7100 | 10 | 60 | 1.968029e+5 |

| | | | | | | | | |
|---------|---------------|----------|---------|--------|-----------|----|----|-------------|
| $m = 2$ | -3.343748e-31 | 0.297990 | 0.00000 | 0.0000 | -74.12700 | 10 | 60 | 4.387772e+3 |
|---------|---------------|----------|---------|--------|-----------|----|----|-------------|

7.3.2、利用解析解求最小 m 值

利用微分方程组的解析解和初始条件构造矢量方程 $A_1 \bar{\alpha} = b_1$ ，此时该方程等式右边的向量 b_1 中包含有一个积分常数 K ，虽然是常数，但其值是未知的。对此式稍作变化：解析解两边同时除以 K ，相当于将解分量 $\alpha_1 \sim \alpha_4$ 同时缩小了 K 倍，则有

$$\begin{aligned} \alpha_1' \ln y + \alpha_2' y - \alpha_3' \ln x - \alpha_{4x}' &= 1 \\ \alpha_5 &= x(1), \alpha_6 = y(1) \end{aligned}$$

对应得向量变为：

$$b_1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ x(1) \\ y(1) \end{bmatrix}_{(2n+2) \times 1}$$

此时亦可以求出不同 m 值时对应得 $|A^T A|$ 、 α 和误差 J ，如表 2。（模拟时为减少误差，采用改进的欧拉方法---预报校正法）

表 2 连续情况下，不同 m 值时对应得 $|A^T A|$ 、 α 和误差 J

| 参 数 组 | 参 数 值 | $ A^T A $ | α_1 | α_2 | α_3 | α_4 | α_5 | α_6 | 模拟误差 |
|-------------|-------------|--------------|------------|------------|--------------|------------|------------|------------|------------------|
| $m = 3$ | | 1.965255e+6 | 2.1707 | -0.21707 | -13.025 | 1.0854 | 10 | 60 | 1.802354e+6 |
| $m = 5$ | | 2.108208e+4 | 2.1706 | -0.217 | -13.025 | 1.0854 | 10 | 60 | 7.678974e+2 |
| $m = 4$ | | 2.504550e-10 | -3.2398 | 2.4833 | -2.5602 | -0.16348 | 10 | 60 | 5.735110e+6 |
| $m = 3$ | | 5.415568e-27 | 9.9162 | -2.9495 | -12.359 | -0.39433 | 10 | 60 | 1.804139e+1 4 |
| $m = 2$ | | 3.503917e-47 | -3.4029 | 60 | -1.1774e-015 | 0 | 10 | 60 | 4.229816e+1 1 |

由上表，行列式 $|A^T A|$ 在 $m=2, 3, 4$ 时均很小，可以看作是零，则此三种情况下原二元线性方程组无解或者解不惟一，从 $m=5$ 开始 $|A^T A|$ 是大于零的数，最

小的观测数据组数必须大于或等于 5。比较 $m=5, 6$ 两种情况的模拟误差知，前者的误差要远小于后者，也即 $m=5$ 时，解的稳定性比较好。由此可以得出结论：连续的情况下至少需要 5 组观测数据才能完全确定参数 α 。

表 2-2 的数据分析还在一定程度上检验了最小二乘法的稳定性。当观测数据组数很小时，如 $m=2, 3$ ，参数解变化剧烈，行列式近似为零，方程组是欠定的。误差也很大；当观测数据增多到一定时，如 $m=5, 6$ ，参数解相差很小，基本趋于稳定了。当然，数据越多，参数解就会越精确。

7.3.3、两种情况的比较

综合上述两者情况可以得出，至少需要 3 组数据才能确定参数 α_i 。图 2 和图 3 是连续情况下用 data1.txt 的前 5 组数据确定参数，然后模拟并与观测数据的比较情况。模拟时采用改进的欧拉方法—预报校正方法。由于数据较少，模拟具有一定的误差，但能够反映出总体的趋势。总的来说离散情况下的模拟效果要比连续情况下稍差。

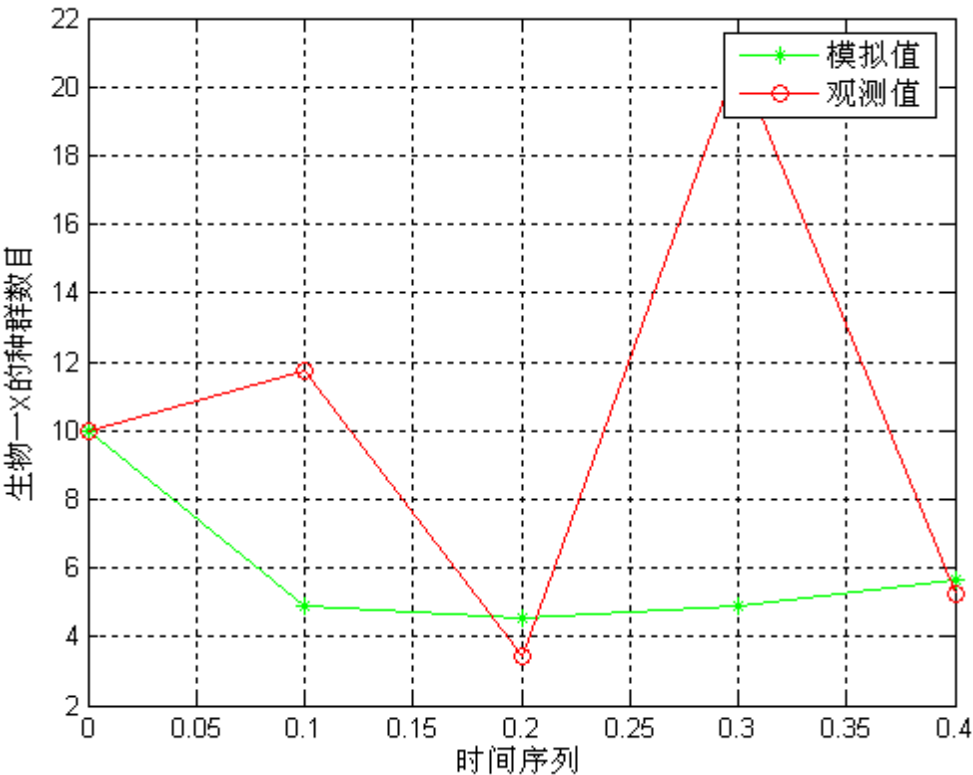


图 2 用 data1.txt 前 5 组观测数据确定的参数进行模拟对比（解析解情况）

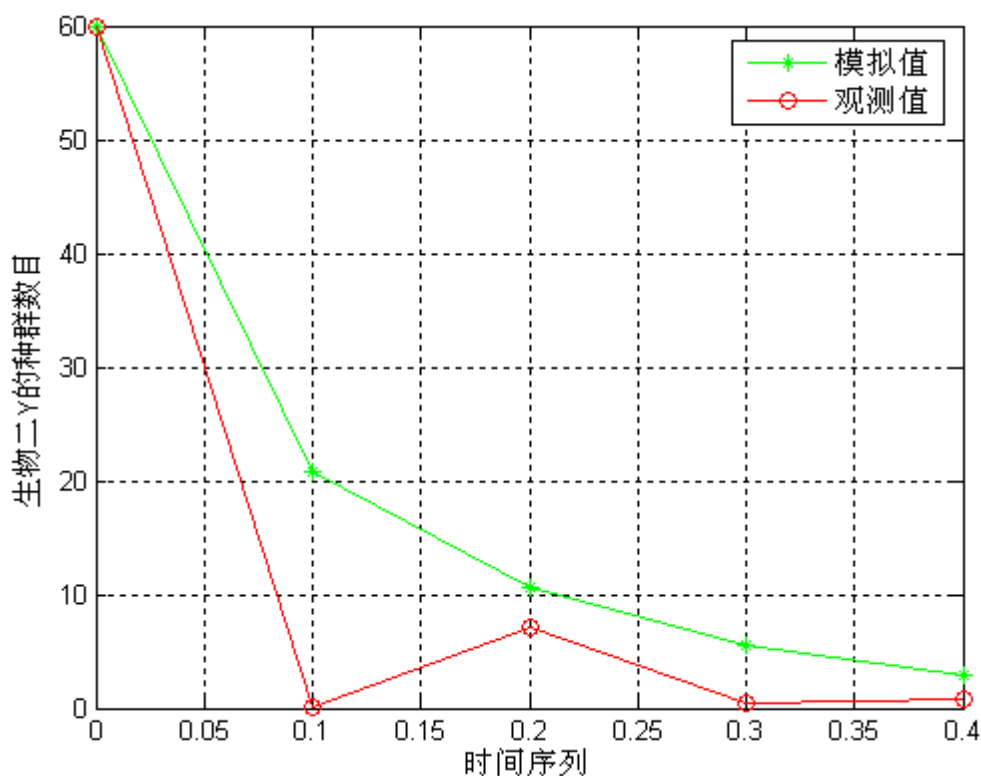


图 3 用 data1.txt 前 5 组观测数据确定的参数进行模拟对比（解析解情况）

八、问题 II -3—高精度差商的最小二乘法模型

在用高精度差商的最小二乘法模型求解的之前,我们还是用问题一的解析解方法试着求解了一下,结果求不出来,但是在理论上还是有指导作用的,在这里也先把它写出来,起到抛砖引玉的作用.

8.1 解析解模型的求解和讨论

8.1.1 解析解模型的求解:

对题目所给的微分方程组模型:

$$\begin{cases} \frac{dx}{dt} = x(t)[\alpha_1 + \alpha_2 y(t)] \\ \frac{dy}{dt} = y(t)[\alpha_3 + \alpha_4 x(t)] \end{cases}$$

在 6.1 节求出其通解为：

$$\alpha_1 \ln y + \alpha_2 y - \alpha_3 \ln x - \alpha_4 x = K \quad (12)$$

其中 K 为任意常数，如果给定了初始条件 (x_0, y_0) ，则 K 为定值。即：

$$K = \alpha_1 \ln y_0 + \alpha_2 y_0 - \alpha_3 \ln x_0 - \alpha_4 x_0$$

由于 K 为任意常数，所以可以将方程(12)两边同除 K ，可得：

$$\frac{\alpha_1}{K} \ln y + \frac{\alpha_2}{K} y - \frac{\alpha_3}{K} \ln x - \frac{\alpha_4}{K} x = 1 \quad (13)$$

将 $\frac{\alpha_1}{K}$, $\frac{\alpha_2}{K}$, $\frac{\alpha_3}{K}$, $\frac{\alpha_4}{K}$ 重新看成新的待求未知数，设为 Z_1, Z_2, Z_3, Z_4 。然后

用 data2 中的数据带入 $\ln y$, y , $\ln x$, x 求出系数阵 A ，这样就是解 $AZ = I$ 的矩阵了。可以求出：

$$Z_1 = 0.14111, Z_2 = -0.0070734, Z_3 = -0.71664, Z_4 = 0.071691$$

代入(13)式，这样就可以求出关于 x 和 y 的关系式：

$$0.14111 \ln y - 0.0070734 y + 0.71664 \ln x - 0.071691 x = 1$$

这就是求出的解析解模型。

8.1.2 解析解模型的讨论：

根据前面求出的解析解模型，我们根据实际给出的 X 值，就可以求出相应的 \hat{Y} 值，将这时的求出的 (X, \hat{Y}) 和实际的 (X, Y) 进行比较，画图如下：

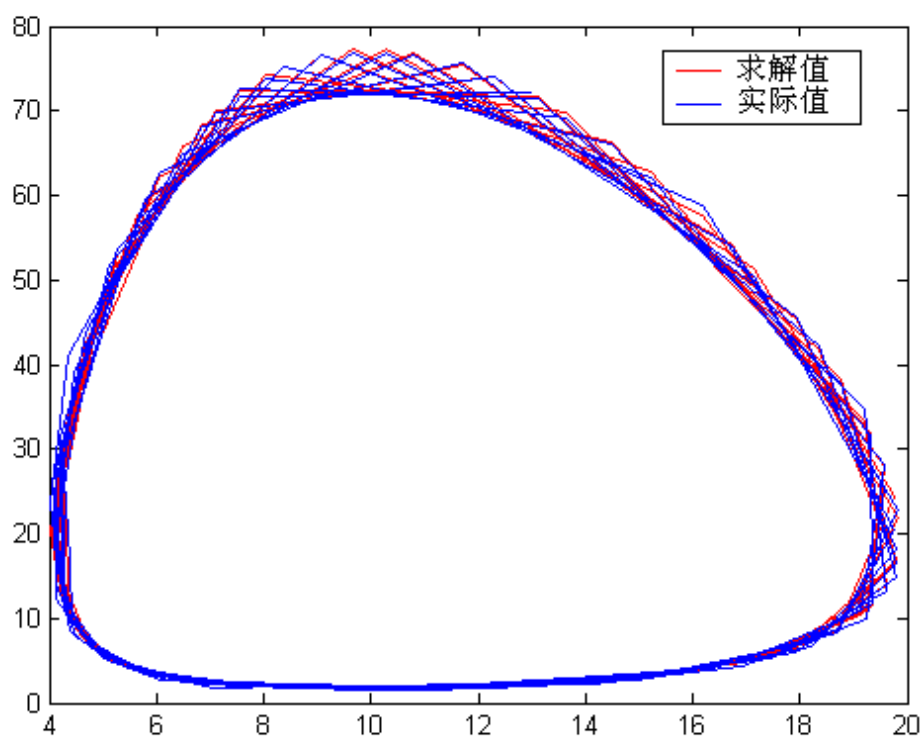


图 4 求出的 (X, \hat{Y}) 和实际的 (X, Y) 比较图

求解值和实际值吻合的很好，可见我们这个解析解的模型还是比较准确的。但是由于 K 的值并不知道,所以此题并不能求出 $\alpha_k (k = 1, 2, 3, 4, 5, 6)$, 但是其在理论上有很好的指导意义！

8.2 考虑了高精度差商的最小二乘法模型的分析 and 求解

问题三在观测资料有误差（时间变量不含有误差）的情况下，同样可以通过最小二乘方法来解决。在此主要将不同于问题二的部分加以讨论。

8.2.1、分析与模型变化

此时观测数据比 data1.txt 增多，可以将问题二的模型基础上，将一阶微商设置成比一阶差商精度更高的差商格式，过程如下：

由泰勒级数：

$$f(x + \Delta x) = f(x) + f'(x)\Delta x + \frac{f''(x)}{2!}(\Delta x)^2 + \frac{f'''(x)}{3!}(\Delta x)^3 + \dots$$

可以写出如下的变形式：

$$\begin{aligned} f(x - \Delta x) &= f(x) - f'(x)\Delta x + \frac{f''(x)}{2!}(\Delta x)^2 - \frac{f'''(x)}{3!}(\Delta x)^3 + \dots \\ f(x + 2\Delta x) &= f(x) + f'(x)2\Delta x + \frac{f''(x)}{2!}(2\Delta x)^2 + \frac{f'''(x)}{3!}(2\Delta x)^3 + \dots \\ f(x - 2\Delta x) &= f(x) - f'(x)2\Delta x + \frac{f''(x)}{2!}(2\Delta x)^2 - \frac{f'''(x)}{3!}(2\Delta x)^3 + \dots \end{aligned}$$

从上面四个展式出发，可以构造出具有 4 阶精度的一阶导数的差分表达式：

$$-f(x + 2\Delta x) + 8f(x + \Delta x) - 8f(x - \Delta x) + f(x - 2\Delta x) = 12\Delta x f'(x)$$

所以得到

$$f'(x) = \frac{-f(x + 2\Delta x) + 8f(x + \Delta x) - 8f(x - \Delta x) + f(x - 2\Delta x)}{12\Delta x}$$

利用上面的结果，则此时离散形式下的矢量方程等式右端的向量可以改写

成：

$$b = \begin{bmatrix} \frac{-x(t_{i+2}) + 8x(t_{i+1}) - 8x(t_{i-1}) + x(t_{i-2}))}{12\Delta t} \\ \frac{-y(t_{i+2}) + 8y(t_{i+1}) - 8y(t_{i-1}) + y(t_{i-2}))}{12\Delta t} \\ x(t_1) \\ y(t_2) \end{bmatrix}_{(2n+2) \times 1}$$

其中， $i = 3, 4, \dots, m-2; n = m-4$;

（说明：该方法要求有 $m > 4$ 组数据，虽然精度很高，但只适合于观测数据

较多的情况，比如 data2.txt 和 data3.txt。）

原微分方程组对应的龙格库塔方法下的差分方程方程组为：

$$\begin{cases} x'(t) = x(t)[\alpha_1 + \alpha_2 y(t)] = f(x, y) \\ y'(t) = y(t)[\alpha_1 + \alpha_2 x(t)] = g(x, y) \end{cases}$$

$$\begin{cases} x_{n+1} = x_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \\ y_{n+1} = y_n + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4) \end{cases}$$

其中，

$$\begin{cases} h = \Delta t \\ k_1 = hf(x_n, y_n), & l_1 = hg(x_n, y_n) \\ k_2 = hf(x_n + \frac{k_1}{2}, y_n + \frac{l_1}{2}) & l_2 = hg(x_n + \frac{k_1}{2}, y_n + \frac{l_1}{2}) \\ k_3 = hf(x_n + \frac{k_2}{2}, y_n + \frac{l_2}{2}) & l_3 = hg(x_n + \frac{k_2}{2}, y_n + \frac{l_2}{2}) \\ k_4 = hf(x_n + k_3, y_n + l_3) & l_4 = hg(x_n + k_3, y_n + l_3) \end{cases}$$

8.2.2、结果讨论

对于 DATA2 数据，由上面的分析,再用最小二乘法编写程序三求出参数 α_i , 为

$$\alpha_1 = 2.0177602612, \alpha_2 = -0.1010220954, \alpha_3 = -9.9028858801$$

$$\alpha_4 = 0.9914774599, \alpha_5 = 12.9622856330, \alpha_6 = 72.1230158334$$

将参数 α_i 带入原方程组，采用龙格库塔方法模拟出的结果来看，模拟值和观测数据吻合的非常好，数据的相关系数达到 99.706%和 99.937%。参考图 5，图 6。

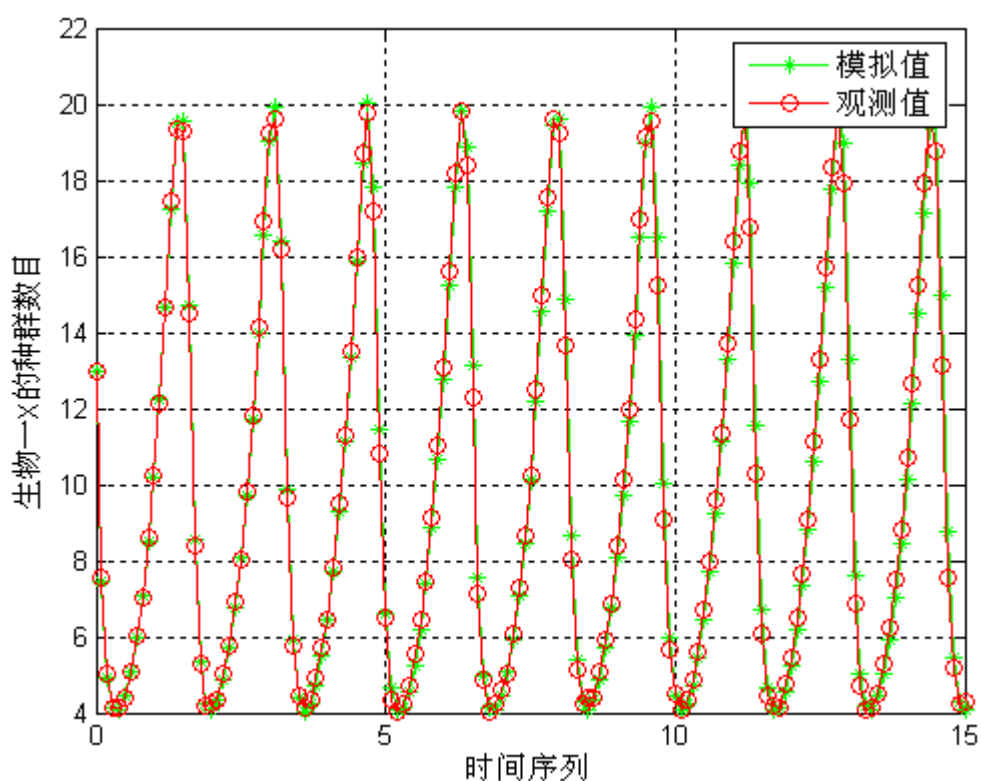


图 5 龙格库塔方法对种群 X 的模拟值和观测值的比较图

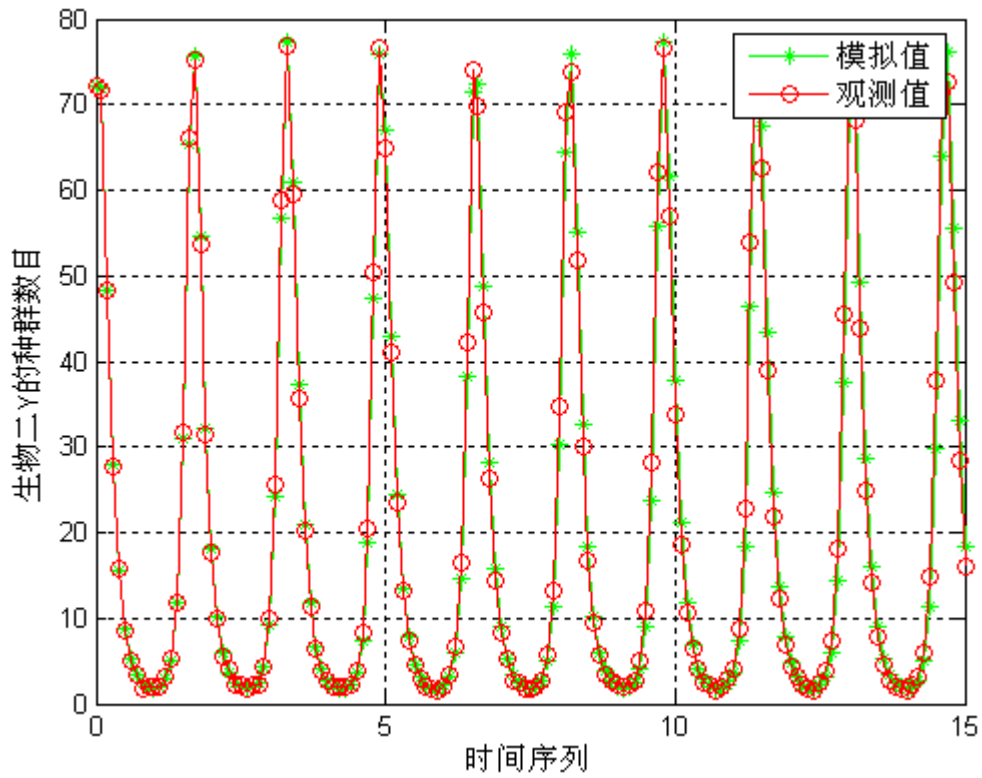


图 6 龙格库塔方法对种群 Y 的模拟值和观测值的图

同样的，对于 DATA3 数据，由上面的分析,再用最小二乘法编写程序三求出参数 α_i ,为

$$\alpha_1 = 2.0162666600, \alpha_2 = -0.1014612382, \alpha_3 = -10.0065548293$$

$$\alpha_4 = 1.0021948539, \alpha_5 = 12.8230578535, \alpha_6 = 73.4072806909$$

将参数 α_i 带入原方程组，采用龙格库塔方法模拟出的结果来看，模拟值和观测数据吻合的非常好，数据的相关系数达到 99.665%和 99.558%。参考图 7，图 8。

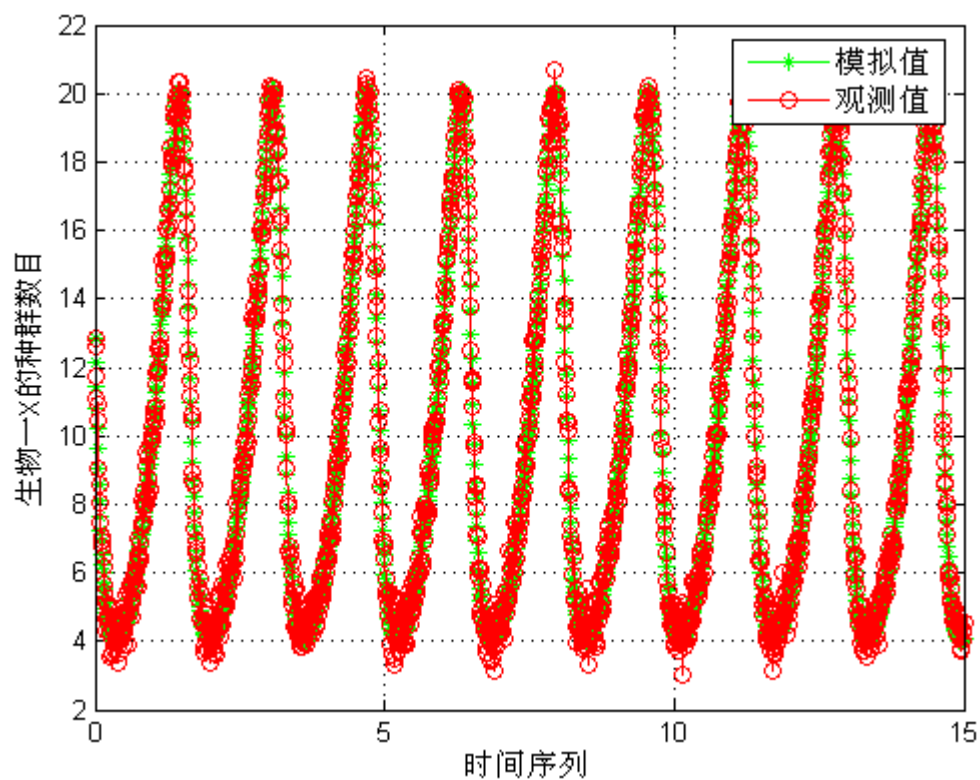


图 7 龙格库塔方法对种群 X 的模拟值和观测值的比较图

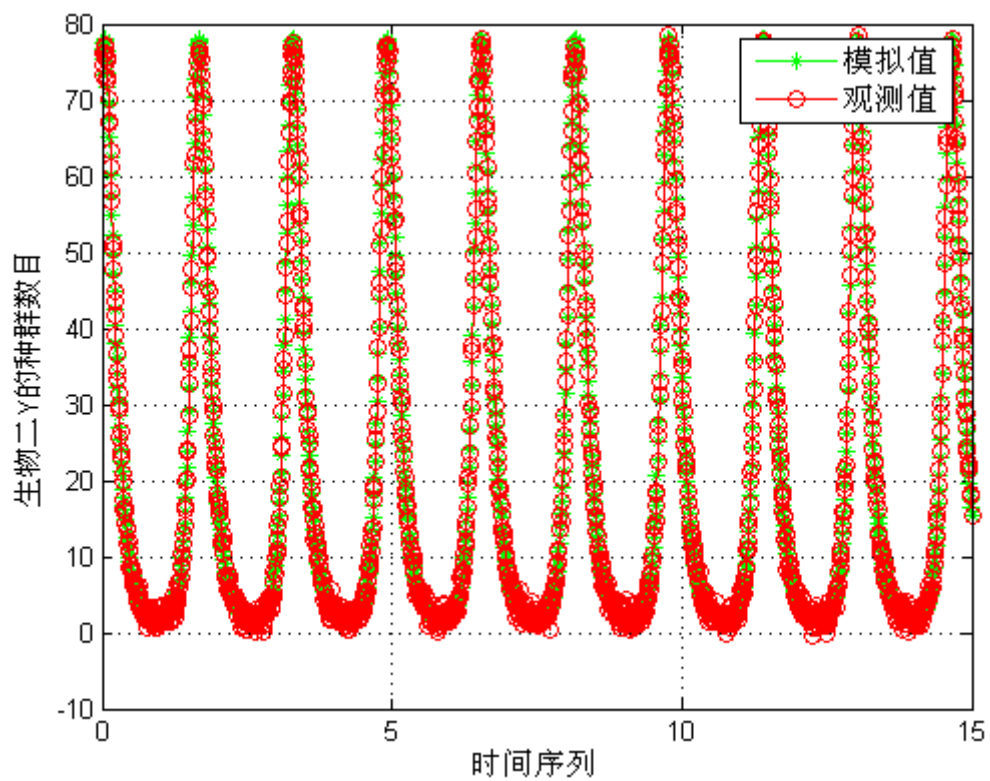


图 8 龙格库塔方法对种群 Y 的模拟值和观测值的图

通过对 DATA2 和 DATA3 的仿真，我们发现用高精度差商的最小二乘法能够仿真真实值相当的好。但是由于题目中说了观测资料的存在误差，所以我们还要对其误差敏感度进行检验。

8.2.3 对模型误差敏感度的检验：

由于问题三中已经明确的说了：在观测资料有误差的情况下，这就是说观测资料并不是真实值，其与真实值之间存在一个随机的误差。这实际上就是要考虑对模型误差敏感度的检验。即对于我们的模型，输入数据的随机误差将会导致模型求解的待定参数发生什么样的变化。

具体来说，要考虑资料误差对本模型的影响，可以通过在资料数据上叠加一个随即的小扰动，即对 data2.txt 和 data3.txt 中除时间数据外（时间无误差）的 x_j ， y_j 添加一个随机扰动量： $(randn - 0.5) * 10^{-N}$ ， $N = 1, 2, \dots$ 。 N 取不同的值代表不同的观测数据误差，计算出不同程度误差时最小二乘意义下的参数 α_k ，并用龙格库塔方法检验了本模型对误差的抗干扰能力。

表 3 是给定 data2.txt 某一量级（精确到小数点后某一位）的扰动后，最小二乘法计算出的参数 α_i 的变化趋势情况。当 N 固定时，由于每次由 $(randn - 0.5) * 10^{-N}$ 提供的扰动量会在一定范围内浮动，为了具有代表意义，我们计算了 50 次的随机扰动参数，并将其平均值作为观测数据误差在 10^{-N} 量级时的最终参数。对于不同的观测误差（ N 取不同的值）， α_i 是不同的。

表 3 DATA2 输入误差对最小二乘模型待定系数的关系表

| 误差量级 系 数 | 10^0 | 10^{-1} | 10^{-2} | 10^{-3} |
|-------------|--------------|--------------|--------------|--------------|
| α_1 | 0.7702891499 | 2.1039745140 | 2.0239582439 | 2.0189241709 |

| | | | | |
|------------|---------------|---------------|---------------|---------------|
| α_2 | -0.0187768069 | -0.1135465991 | -0.1019492709 | -0.1011432869 |
| α_3 | -1.2771668557 | -8.9348012872 | -9.8261194113 | -9.8938528635 |
| α_4 | 0.3939214717 | 1.0289890673 | 0.9961024487 | 0.9918026097 |
| α_5 | -1.0762226493 | 11.2596012863 | 12.8740250391 | 12.9470374892 |
| α_6 | 62.1474009199 | 71.2323393116 | 71.9891819498 | 72.1180942750 |

| 误差量级 系数 | 10^{-4} | 10^{-5} | 10^{-6} | 无误差 |
|------------|---------------|---------------|---------------|---------------|
| α_1 | 2.0178532851 | 2.0177678388 | 2.0177612497 | 2.0177602612 |
| α_2 | -0.1010329755 | -0.1010231595 | -0.1010222174 | -0.1010220954 |
| α_3 | -9.9021324132 | -9.9028097005 | -9.9028756171 | -9.9028858801 |
| α_4 | 0.9915289954 | 0.9914820613 | 0.9914778036 | 0.9914774599 |
| α_5 | 12.9619483272 | 12.9620922483 | 12.9622713953 | 12.9622856330 |
| α_6 | 72.1215504025 | 72.1228697125 | 72.1229974818 | 72.1230158334 |

同样对于 DATA3 数据,我们将输入的随机误差量级分别从 10^0 到 10^{-6} 进行讨论,可以看出其导致的模型求解的待定系数变化如下表所示:

表 4 DATA3 输入误差对最小二乘模型待定系数的关系表

| 误差量级 系数 | 10^0 | 10^{-1} | 10^{-2} | 10^{-3} |
|------------|---------------|---------------|---------------|---------------|
| α_1 | 0.6211920781 | 2.1079306724 | 2.0255649708 | 2.0173438968 |
| α_2 | -0.0171934995 | -0.1136828939 | -0.1026290525 | -0.1015820410 |
| α_3 | -0.6570870256 | -8.9091525318 | -9.9149062105 | -9.9973244883 |

| | | | | |
|------------|---------------|---------------|---------------|---------------|
| α_4 | 0.4654530800 | 1.0207710213 | 1.0060525669 | 1.0025543511 |
| α_5 | -2.9319960450 | 10.7801430453 | 12.7143640674 | 12.8113231987 |
| α_6 | 62.0669372598 | 72.1369354706 | 73.2654843785 | 73.3973888601 |

| 误差量级 系数 | 10^{-4} | 10^{-5} | 10^{-6} | 无误差 |
|------------|----------------|----------------|----------------|----------------|
| α_1 | 2.0163583406 | 2.0162771045 | 2.0162677273 | 2.0162666600 |
| α_2 | -0.1014721858 | -0.1014624278 | -0.1014613596 | -0.1014612382 |
| α_3 | -10.0056597473 | -10.0064616345 | -10.0065458302 | -10.0065548293 |
| α_4 | 1.0022336933 | 1.0021981820 | 1.0021952390 | 1.0021948539 |
| α_5 | 12.8219115278 | 12.8229487752 | 12.8230444174 | 12.8230578535 |
| α_6 | 73.4058579038 | 73.4071849764 | 73.4072731058 | 73.4072806909 |

从上面的两个表上就可以清楚地看出，当输入误差的量级小于 10^{-2} 的时候，我们模型的待定系数都还是变化不大的，比如说对于 DATA2 数据，当误差量级为 10^{-2} 时，数据的相关系数仍然能够达到 98.53%和 97.95%。如图 9，图 10。

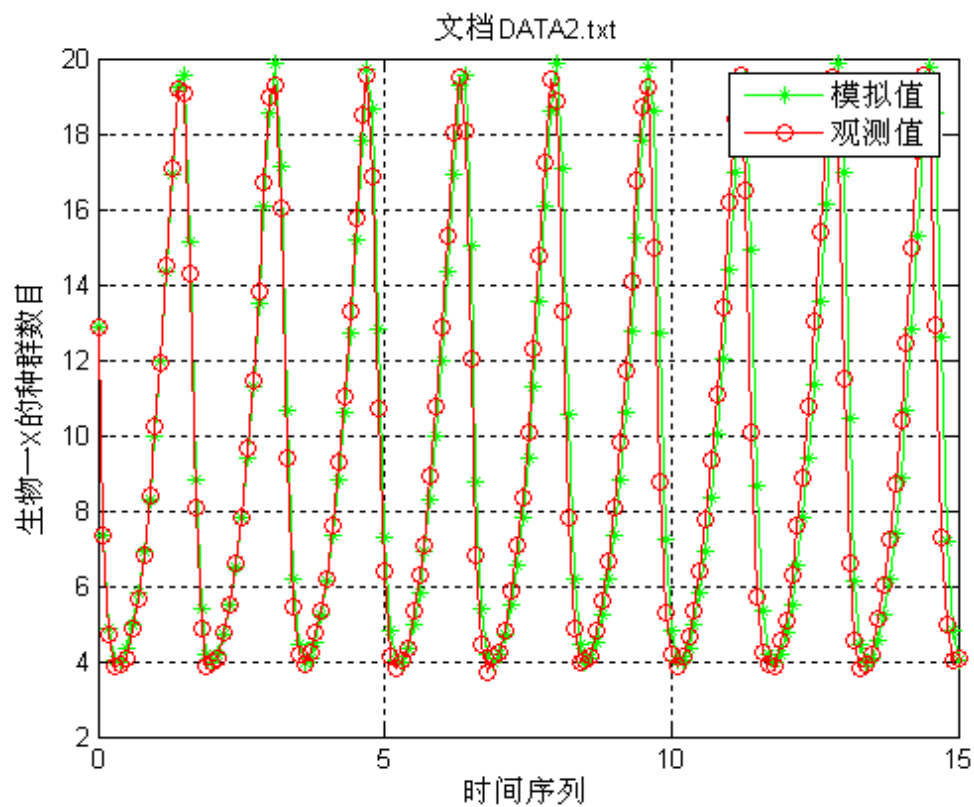


图9 当误差量级为 10^{-2} 时, 龙格库塔方法对种群 X 的模拟图

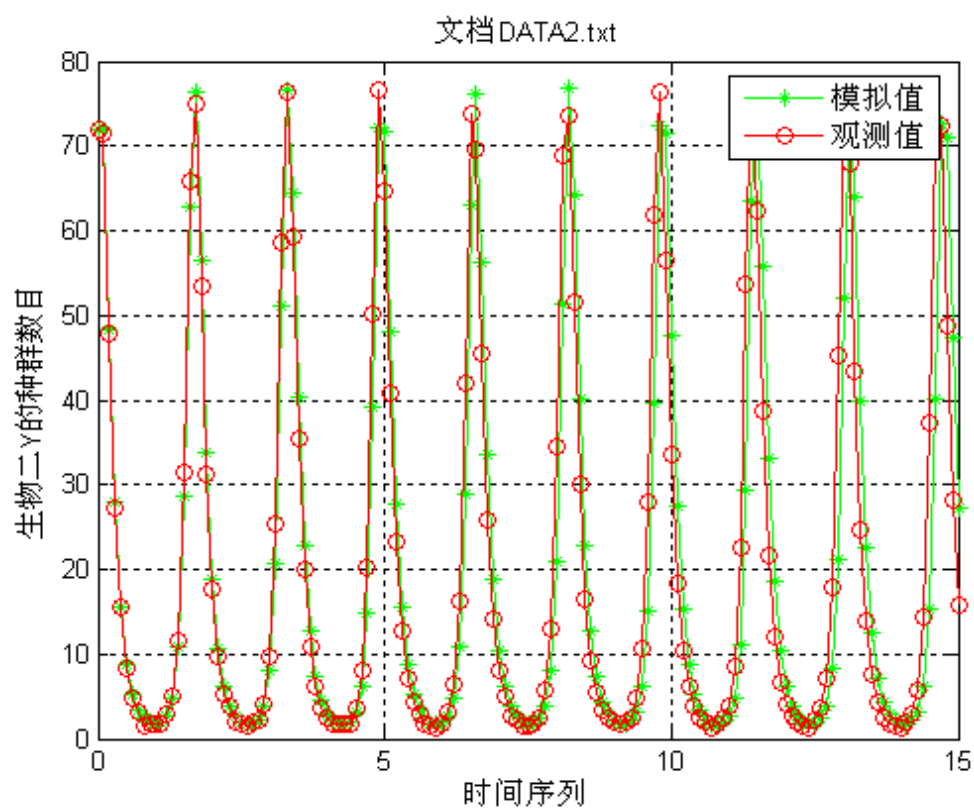


图10 当误差量级为 10^{-2} 时, 龙格库塔方法对种群 Y 的模拟图

而且是输入误差的量级越小，我们的待定系数的变化也是越小，基本改变量级等于输入误差的量级。说明我们的这个模型对误差并不是很敏感，是比较稳定的。

但是当输入误差的量级大于 10^{-2} 的时候，待定系数 α 的变化就更大了。特别是量级大到 10^0 的时候，待定系数就已经完全改变了。可见此模型还可以继续改进，来提高误差敏感性，变得更稳定。

这样我们就需要在最小二乘模型的基础上引入一个更好的改进模型——变分伴随同化模型。

九、问题 II -3—变分伴随同化模型

引入变分伴随同化模型，除了希望得到比最小二乘法模型更优的参数值外，更重要的希望能找到一个更稳定抗误差敏感的模型，同时可以对变参数进行反演。而变分伴随方法正好有这些优点。

9.1 问题的分析

通过分析题中给出捕食者和被捕食者生态系统种群数目变化规律的数学模型(如下:3-1 式)可以看出,该模型的解,即 t 时刻 A 和 B 生物种群的数目 $x(t), y(t)$ 是由模型中的参数唯一确定的。

$$\begin{cases} \frac{dx}{dt} = \alpha_1 x(t) + \alpha_2 x(t)y(t) \\ \frac{dy}{dt} = \alpha_3 y(t) + \alpha_4 x(t)y(t) \\ x(t_0) = \alpha_5 \\ y(t_0) = \alpha_6 \end{cases}, (t_0 \leq t \leq T) \quad (3-1)$$

其中 $\alpha_k (1 \leq k \leq 6)$ 为参数。。

首先介绍连续系统（或假定是连续的观测资料）中反演模型参数

$\alpha_j, (j=1, \dots, 6)$ 的数学方法，先建立如下形式的目标泛函：

$$J(\alpha) = \frac{1}{2} \int_{t_0}^T (x(t) - x^{obs}(t))^2 dt + \frac{1}{2} \int_{t_0}^T w(y(t) - y^{obs}(t))^2 dt$$

其中 $x^{obs}(t), y^{obs}(t)$ 分别代表 t 时刻观测到的 A 和 B 生物种群的数量， $\alpha = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \alpha_4 \ \alpha_5 \ \alpha_6]$ ， w 为权重，其作用是使 $x(t)$ 和 $y(t)$ 的数量级是相同的。目标泛函 J 实际上代表了 t_0 到 T 时刻模型解与观测值的总偏差。模型中参数的最优解也就是使目标泛函达到最小。这是一个泛函的条件极值问题。求解泛函的极小值需要采用最优化算法，其基本思想如下：给定 α 一个初值 α^0 ，

按照下式进行迭代：

$$\alpha^{k+1} = \alpha^k + d^k \rho^k$$

其中 d^k 搜索方向， ρ^k 为步长。

每次迭代得到新的 α 值后判断 $J(\alpha^{k+1}) < J(\alpha^k)$ 是否成立，不成立则通过调整步长，使得 α 向着目标泛函减小的方向进行迭代。当 $J(\alpha^{k+1})$ 小于一个给定的小量 ε 时退出迭代过程，此时得到的 α^{k+1} 即可以认为是参数 α 的最优解。

在变分伴随方法中，采用 d^k 为泛函的负梯度方向：

$$d_k = -(\nabla_{\alpha} J)(\alpha^k)$$

于是所求泛函的梯度。

结合上面的模型，我们采用伴随方法是问题的关键所在，求解上面建立的目标泛函的梯度。所谓变分伴随同化方法是变分结合伴随理论而提出的一种方法，具体实施方案如下：

Step1: 令

$$\begin{cases} \tilde{x} = x + \hat{x}\theta \\ \tilde{y} = y + \hat{y}\theta \\ \tilde{\alpha}_i = \alpha_i + \hat{\alpha}_i\theta, (i=1,2,\dots,6) \end{cases} \quad (3-2)$$

其中 θ 代表扰动的大小， $\hat{x}, \hat{y}, \hat{\alpha}_i (i=1,2,\dots,6)$ 代表扰动的方向。用 (\tilde{x}, \tilde{y}) 表示 (3-1)

进行扰动后方程的解，此时 (\tilde{x}, \tilde{y}) 满足：

$$\begin{cases} \dot{\tilde{x}} = \tilde{\alpha}_1 \tilde{x} + \tilde{\alpha}_2 \tilde{x} \tilde{y} \\ \dot{\tilde{y}} = \tilde{\alpha}_3 \tilde{y} + \tilde{\alpha}_4 \tilde{x} \tilde{y} \\ \tilde{x}|_{t=t_0} = \tilde{\alpha}_5 \\ \tilde{y}|_{t=t_0} = \tilde{\alpha}_6 \end{cases} \quad (3-3)$$

(3-3) 式减去 (3-2) 再除以 θ ，并令 $\theta \rightarrow 0$ ，记 $(\hat{x}, \hat{y}) = \lim_{\theta \rightarrow 0} \frac{(\tilde{x}, \tilde{y}) - (x, y)}{\theta}$ ，将 (3-2) 式代入 (3-3) 式中，则有：

$$\begin{cases} \dot{\hat{x}} = \alpha_1 \hat{x} + \hat{\alpha}_1 x + \alpha_2 \hat{x} y + \alpha_2 x \hat{y} + \hat{\alpha}_2 xy \\ \dot{\hat{y}} = \alpha_3 \hat{y} + \hat{\alpha}_3 y + \alpha_4 \hat{x} y + \alpha_4 x \hat{y} + \hat{\alpha}_4 xy \\ \hat{x}|_{t=t_0} = \hat{\alpha}_5 \\ \hat{y}|_{t=t_0} = \hat{\alpha}_6 \end{cases} \quad (3-4)$$

Step2: 求泛函的梯度，记 $J'(\alpha; \hat{\alpha})$ 为 J 在 α 沿着 $\hat{\alpha}$ 的方向导数，即，

$$\begin{aligned} J'(\alpha; \hat{\alpha}) &= \lim_{\theta \rightarrow 0} \frac{J(\tilde{\alpha}) - J(\alpha)}{\theta}, \text{ 则利用 } J \text{ 的定义及 } J \text{ 的表达式，我们得出：} \\ &J[\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6; \alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, \alpha_6] \\ &= \nabla_{\alpha_1} J \cdot \hat{\alpha}_1 + \nabla_{\alpha_2} J \cdot \hat{\alpha}_2 + \dots + \nabla_{\alpha_6} J \cdot \hat{\alpha}_6 \\ &= \int_{t_0}^T (x - x^{obs}) \hat{x} dt + \int_{t_0}^T w(y - y^{obs}) \hat{y} dt \end{aligned} \quad (3-5)$$

Step3: 引入伴随模式方法，引入伴随量 (P, Q) ，用 P, Q 分别乘 (3-4) 式第一、二式，相加，再从 $t_0 \rightarrow T$ 积分，有：

$$\begin{aligned} \int_{t_0}^T (\dot{\hat{x}} P + \dot{\hat{y}} Q) dt &= \int_{t_0}^T (\alpha_1 \hat{x} + \hat{\alpha}_1 x + \alpha_2 \hat{x} y + \alpha_2 x \hat{y} + \hat{\alpha}_2 xy) P dt \\ &\quad + \int_{t_0}^T (\alpha_3 \hat{y} + \hat{\alpha}_3 y + \alpha_4 \hat{x} y + \alpha_4 x \hat{y} + \hat{\alpha}_4 xy) Q dt \end{aligned} \quad (3-6)$$

(3-6) 式左边利用分部积分的方法可得：

$$\text{左边} = \int_{t_0}^T -\left(\frac{dP}{dt} \hat{x} + \frac{dQ}{dt} \hat{y}\right) dt + \hat{x} P|_{t_0}^T + \hat{y} Q|_{t_0}^T$$

利用 $\hat{x}|_{t=t_0} = \hat{\alpha}_5, \hat{y}|_{t=t_0} = \hat{\alpha}_6$ ，上式化为：

$$\text{左边} = \int_{t_0}^T -\left(\frac{dP}{dt} \hat{x} + \frac{dQ}{dt} \hat{y}\right) dt - \hat{\alpha}_5 P(t_0) - \hat{\alpha}_6 Q(t_0) + \hat{x}(T) P(T) + \hat{y}(T) Q(T)$$

(3-6) 式右边可以化为：

$$\begin{aligned}
\text{右边} = & \int_{t_0}^T (\alpha_1 P + \alpha_2 yP + \alpha_4 yQ) \hat{x} dt + \int_{t_0}^T (\alpha_3 Q + \alpha_4 xQ + \alpha_2 xP) \hat{y} dt \\
& + \hat{\alpha}_1 \int_{t_0}^T xP dt + \hat{\alpha}_2 \int_{t_0}^T xyP dt + \hat{\alpha}_3 \int_{t_0}^T yQ dt + \hat{\alpha}_4 \int_{t_0}^T xyQ dt
\end{aligned}$$

从而

$$\begin{aligned}
& \int_{t_0}^T \left(-\frac{dP}{dt} - \alpha_1 P - \alpha_2 yP - \alpha_4 yQ \right) \hat{x} dt + \int_{t_0}^T \left(-\frac{dQ}{dt} - \alpha_3 Q - \alpha_4 xQ - \alpha_2 xP \right) \hat{y} dt - \hat{\alpha}_5 P(t_0) + \hat{\alpha}_6 Q(t_0) \\
& + \hat{x}(T)P(T) + \hat{y}(T)Q(T) = \hat{\alpha}_1 \int_{t_0}^T xP dt + \hat{\alpha}_2 \int_{t_0}^T xyP dt + \hat{\alpha}_3 \int_{t_0}^T yQ dt + \hat{\alpha}_4 \int_{t_0}^T xyQ dt
\end{aligned}$$

由此可以得到伴随模式:

$$\begin{cases} -\frac{dP}{dt} = \alpha_1 P + \alpha_2 yP + \alpha_4 yQ + (x - x^{obs}) \\ -\frac{dQ}{dt} = \alpha_3 Q + \alpha_4 xQ + \alpha_2 xP + w(y - y^{obs}) \\ P|_{t=T} = 0 \\ Q|_{t=T} = 0 \end{cases} \quad (3-7)$$

$$\begin{aligned}
& \int_{t_0}^T (x - x^{obs}) \hat{x} dt + \int_{t_0}^T w(y - y^{obs}) \hat{y} dt - \hat{\alpha}_5 P(t_0) - \hat{\alpha}_6 Q(t_0) \\
& = \hat{\alpha}_1 \int_{t_0}^T xP dt + \hat{\alpha}_2 \int_{t_0}^T xyP dt + \hat{\alpha}_3 \int_{t_0}^T yQ dt + \hat{\alpha}_4 \int_{t_0}^T xyQ dt
\end{aligned} \quad (3-8)$$

对比(3-5)和(3-7)式可得:

$$\begin{cases} \nabla_{\alpha_1} J = \int_{t_0}^T xP dt, \nabla_{\alpha_2} J = \int_{t_0}^T xyP dt, \nabla_{\alpha_3} J = \int_{t_0}^T yQ dt \\ \nabla_{\alpha_4} J = \int_{t_0}^T xyQ dt, \nabla_{\alpha_5} J = P(t_0), \nabla_{\alpha_6} J = Q(t_0) \end{cases} \quad (3-9)$$

通过(3-8)式计算目标泛函的梯度, 在此基础上采用最优化方法来反演(3-1)式中参数的最优解。这里我们采用变分伴随同化的方法, 其计算流程见图 11 :

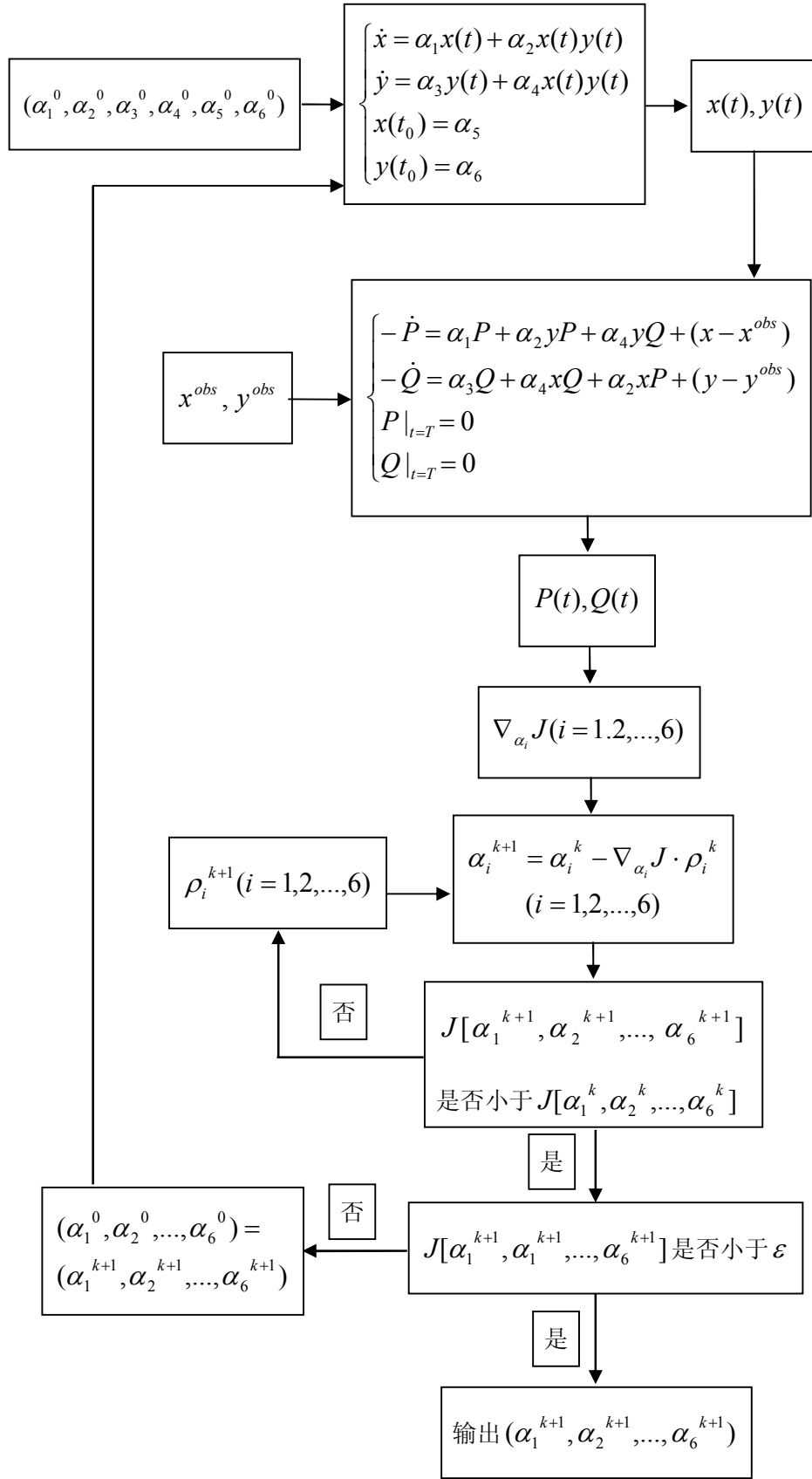


图 11 流程图

9.2 问题的求解：

采用上述变分同化的方法，通过使用 Matlab 和 Fortran 工具，我们对问题二的第三问进行求解。求解过程中采用最小二乘法模型计算得到的 $\alpha_i(i=1,2,...,6)$ 作为初值，得到优化后 $\alpha_i(i=1,2,...,6)$ ，下面分别对 data2 和 data3 的结果进行分析。

9.2.1 data2.txt 的分析结果

表 5 给出了采用变分伴随同化方法对参数进行优化前后的参数，其中优化前的参数为最小二乘法模型计算得到的参数值。从表 6 中可以看出，优化后相关系数有所提高，平均绝对误差明显减小。图 12 和图 13 分别给出了 A 和 B 种群数目的观测值、未优化参数计算的仿真值和优化后参数计算的仿真值。从图中可以看出，虽然优化前后的仿真值都能够很好地反映种群数目的变化趋势，但是优化后结果明显更接近观测值，这也从直观上解释了优化前后相关系数并没有很大提高，而平均绝对误差明显减小。从图 14 上则可以看出，优化后仿真结果的误差比较小.对 A 种群而言，误差在[-0.5, 0.5] 范围内，B 种群，误差在[-0.3, 0.3] 范围内。从相空间平面图上也可以看出，仿真结果与观测值相当吻合。

表 5 优化前后参数对比

| 参数 | 优化前 | 优化后 |
|------------|------------------|------------------|
| α_1 | 2. 017760276794 | 2. 020061124939 |
| α_2 | -0. 101022094488 | -0. 101120996765 |
| α_3 | -9. 902885437012 | -9. 914168816403 |
| α_4 | 0. 991477489471 | 0. 992633766537 |
| α_5 | 12. 962285995483 | 12. 962285996096 |
| α_6 | 72. 123016357422 | 72. 123016357415 |

表 6 优化前后仿真值与观测值的相关系数和平均绝对误差

| 相关系数 | 优化前 | 优化后 |
|--------|-----------------|-----------------|
| A种群 | 0. 998509912713 | 0. 999840119326 |
| B种群 | 0. 998248812921 | 0. 999962731039 |
| 平均绝对误差 | 优化前 | 优化后 |
| A种群 | 0. 207352923138 | 0. 081746791397 |
| B 种群 | 0. 881013008109 | 0. 173674550580 |

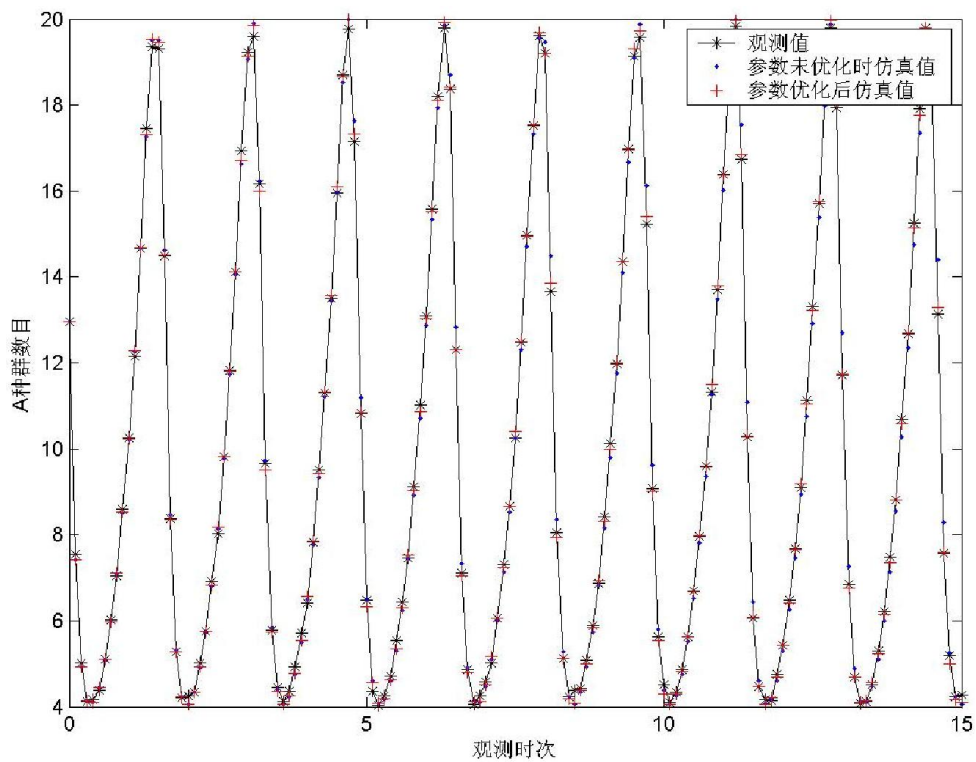


图 12 A 种群数目的观测值、未优化参数计算的仿真值、优化后参数计算的仿真值比较图

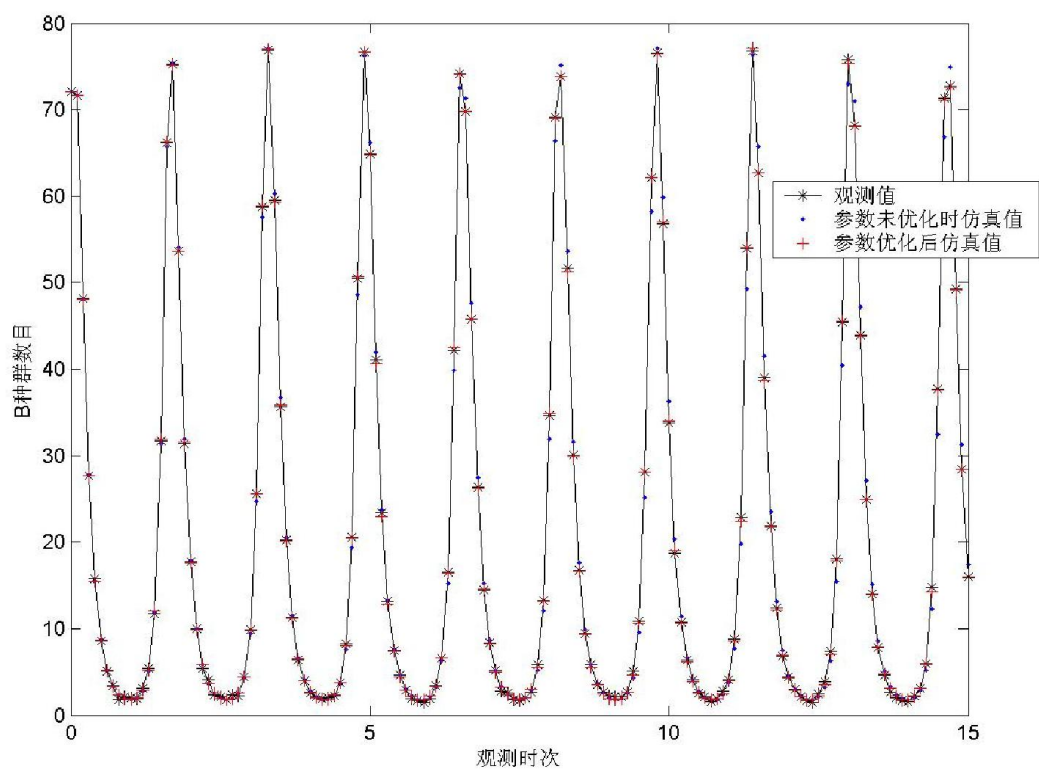


图 13 B 种群数目的观测值、未优化参数计算的仿真值、优化后参数计算的仿真值比较图

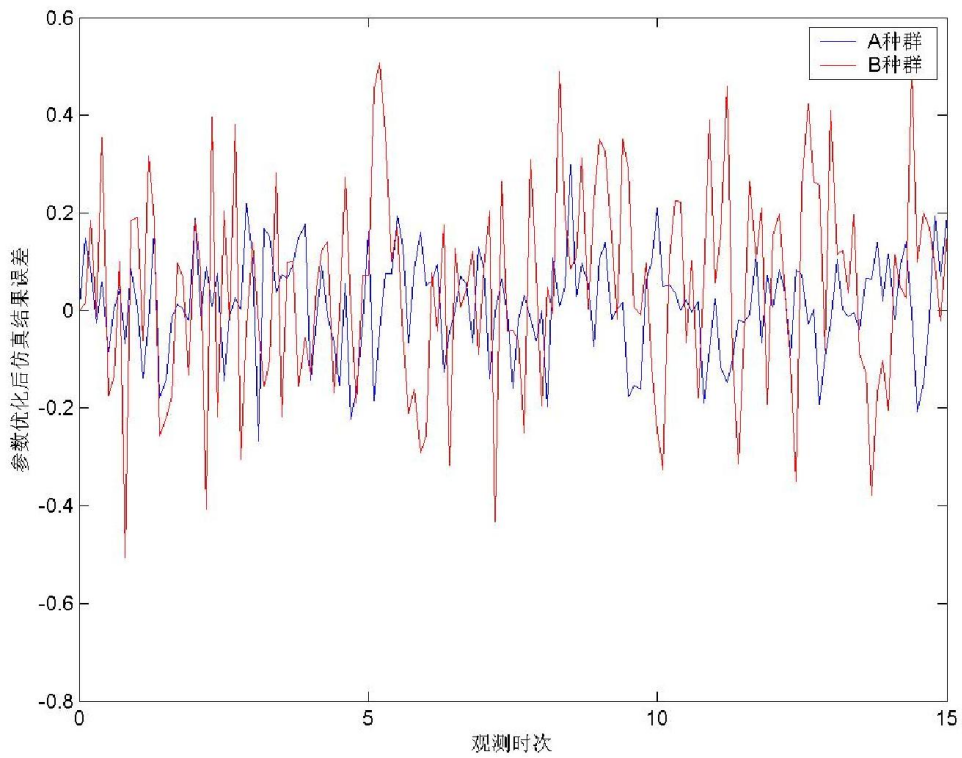


图 14 参数优化后仿真结果误差

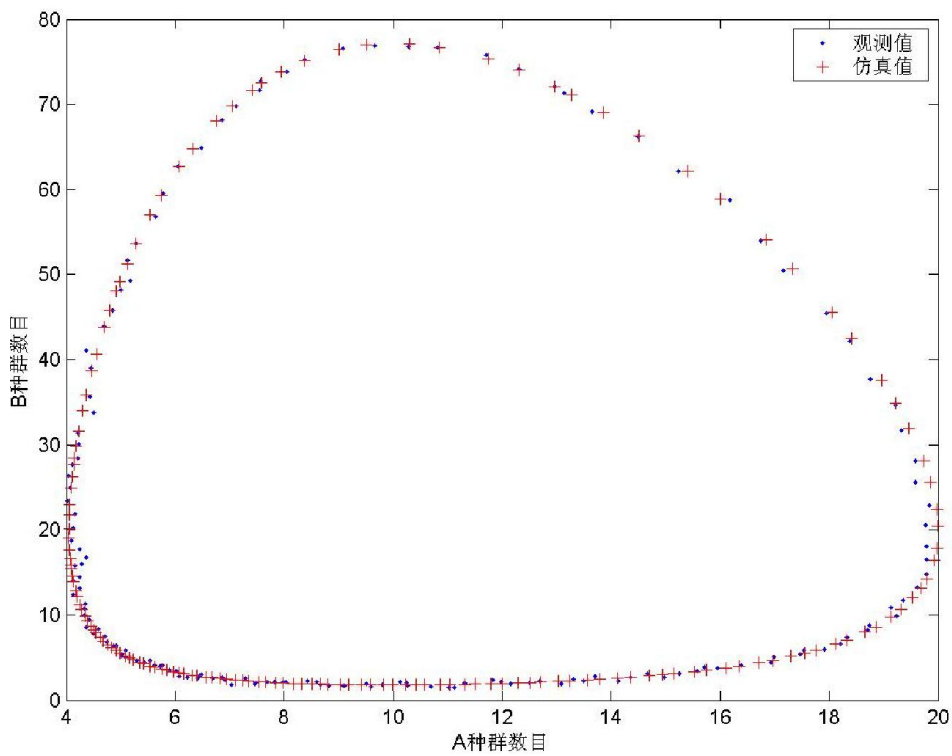


图 15 相空间图

9.2.2 data3.txt 的分析结果

从表 8 中可以看出，采用变分伴随同化方法对参数优化起的效果很小，但还是起到了一定的优化作用。表 7 和表 8 还是优化前的最小二乘法和优化后的参数和相关系数的比较。还可以画出优化后的仿真值和观测值之间的比较图。

表 7 优化前后参数对比

| 参数 | 优化前 | 优化后 |
|------------|------------------|------------------|
| α_1 | 2.016266660000 | 2.016266584396 |
| α_2 | -0.101461238200 | -0.101461239159 |
| α_3 | -10.006554829300 | -10.006554603577 |
| α_4 | 1.002194853900 | 1.002194881439 |
| α_5 | 12.823057853500 | 12.823058128357 |
| α_6 | 73.407280690900 | 73.407279968262 |

表 8 优化前后仿真值与观测值的相关系数和平均绝对误差

| 相关系数 | 优化前 | 优化后 |
|--------|----------------|----------------|
| A种群 | 0.997061902801 | 0.997061902875 |
| B种群 | 0.999373126745 | 0.999373126548 |
| | | |
| 平均绝对误差 | 优化前 | 优化后 |
| A种群 | 0.330362623259 | 0.330362616202 |
| B种群 | 0.704997616206 | 0.704997735823 |

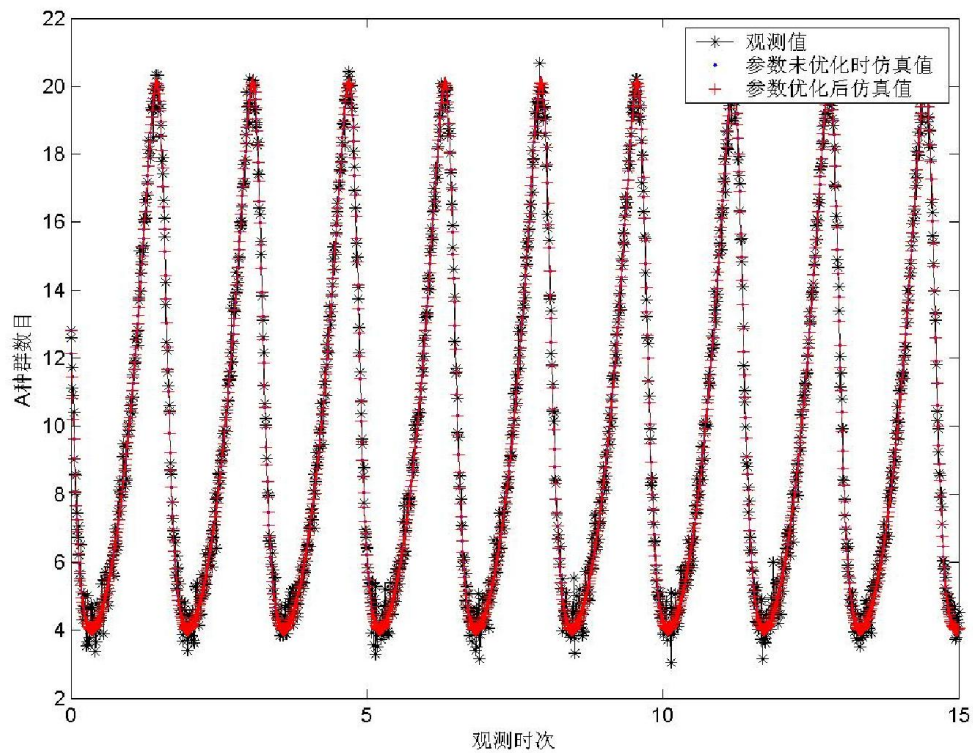


图 16 A 种群数目的观测值、未优化参数计算的仿真值、优化后参数计算的仿真值比较图

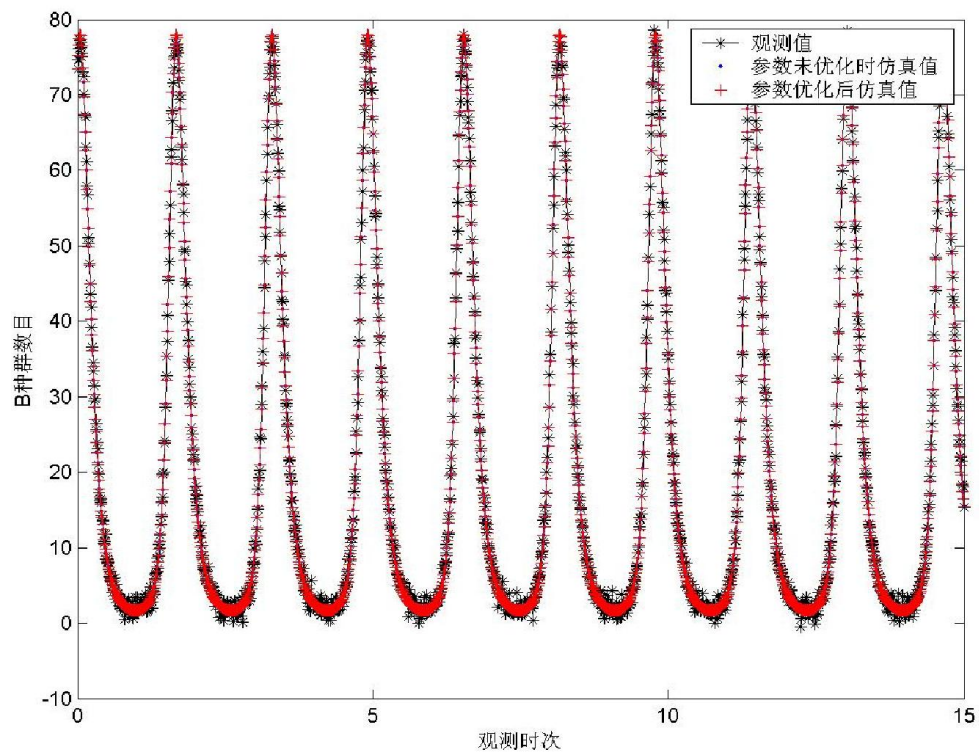


图 17 B 种群数目的观测值、未优化参数计算的仿真值、优化后参数计算的仿真值比较图

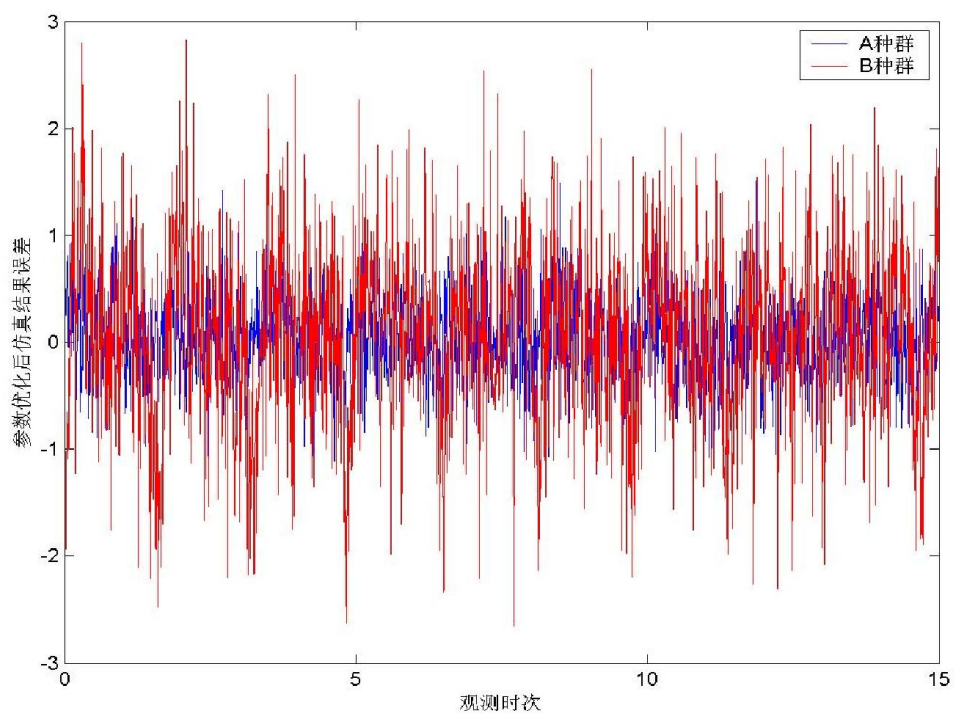


图 18 参数优化后仿真结果误差

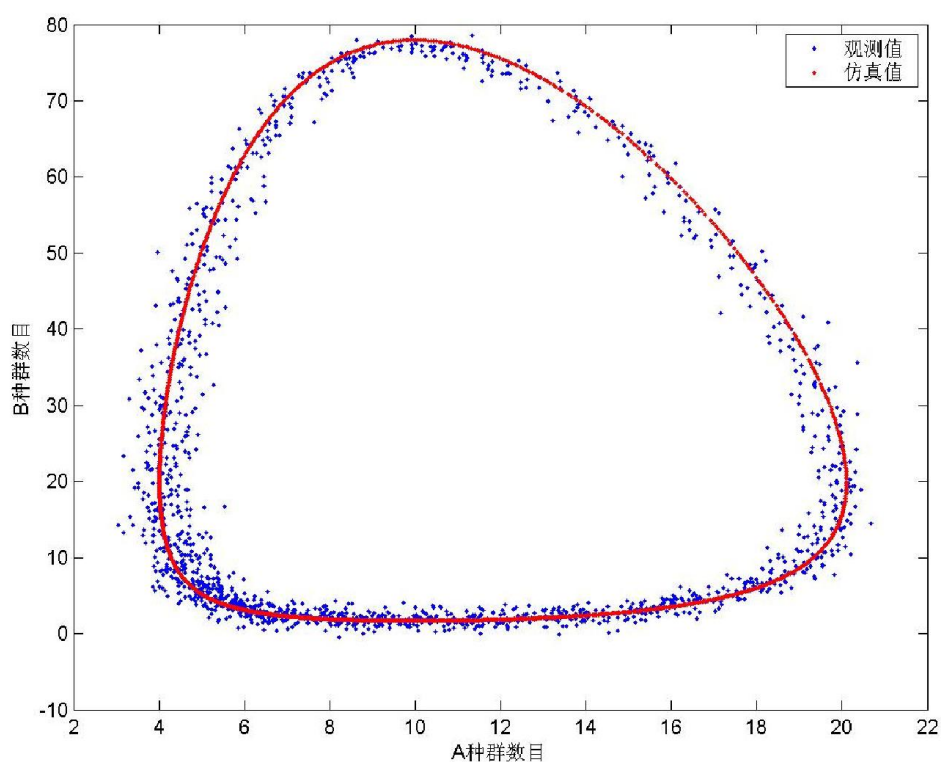


图 19 相空间图

9.3 对模型误差敏感度的检验：

与前面的 8.3.3 一样，我们也要检验变分同化模型的误差敏感度，并且说明其比最小二乘法更稳定，更有优势。我们引入变化伴随同化的最主要目的就是在模型稳定度上改进。

同样，对于变分同化模型，对于 DATA2 数据，我们也将输入的随机误差量级分别从 10^0 到 10^{-6} 进行讨论，可以看出其导致的模型求解的待定系数变化如下表所示：

表 9 DATA2 输入误差对变分同化模型待定系数的关系表

| 误差量级 系数 | 10^0 | 10^{-1} | 10^{-2} | 10^{-3} |
|------------|----------------|----------------|----------------|----------------|
| α_1 | 2.10822326883 | 2.024758492 | 2.02628201820 | 2.02123404906 |
| α_2 | -0.06833135842 | -0.10187198243 | -0.10162138562 | -0.10103585349 |
| α_3 | -8.93704825489 | -9.88904757736 | -9.83717490352 | -9.90505186002 |

| | | | | |
|------------|---------------|----------------|----------------|----------------|
| α_4 | 1.05846424301 | 0.99539776201 | 0.99786693277 | 0.99328570068 |
| α_5 | 1.0709633290 | 11.25960383691 | 12.87402536170 | 12.94703770523 |
| α_6 | 63.187792347 | 71.23233792746 | 71.98918151838 | 72.11809539786 |

| 误差量级 系数 | 10^{-4} | 10^{-5} | 10^{-6} | 无误差 |
|------------|----------------|----------------|----------------|----------------|
| α_1 | 2.02015561957 | 2.02006885991 | 2.02006209793 | 2.02006112493 |
| α_2 | -0.10109388569 | -0.10111944372 | -0.10112062312 | -0.10112099676 |
| α_3 | -9.91340075927 | -9.91409243878 | -9.91415908969 | -9.91416881640 |
| α_4 | 0.992745974848 | 0.99264254534 | 0.99263485250 | 0.99263376653 |
| α_5 | 12.96194839681 | 12.96209240030 | 12.96227169100 | 12.96228599609 |
| α_6 | 72.12155151365 | 72.12287139891 | 72.12300109862 | 72.12301635741 |

同样对于 DATA3 数据,我们将输入的随机误差量级分别从 10^0 到 10^{-6} 进行讨论,可以看出其导致的模型求解的待定系数变化如下表所示:

表 10 DATA3 输入误差对变分同化模型待定系数的关系表

| 误差量级 系数 | 10^0 | 10^{-1} | 10^{-2} | 10^{-3} |
|------------|----------------|----------------|----------------|----------------|
| α_1 | 2.1408266702 | 2.0284681900 | 2.02556575881 | 2.01734399795 |
| α_2 | -0.12547743970 | -0.10399120042 | -0.10260341374 | -0.10158204287 |
| α_3 | -9.89544287435 | -9.98277609122 | -9.91486480809 | -9.99732494354 |
| α_4 | 1.02773219840 | 1.00875310997 | 1.00623232218 | 1.00255429744 |
| α_5 | 4.2533790342 | 10.7190081124 | 12.71436405192 | 12.81132316589 |

| | | | | |
|------------|--------------|--------------|----------------|----------------|
| α_6 | 70.915009543 | 72.098964350 | 73.26548767089 | 73.39739227294 |
|------------|--------------|--------------|----------------|----------------|

| 误差量级 系数 | 10^{-4} | 10^{-5} | 10^{-6} | 无误差 |
|------------|-----------------|-----------------|----------------|----------------|
| α_1 | 2.01635837554 | 2.016267776489 | 2.01626777649 | 2.01626658439 |
| α_2 | -0.10147218406 | -0.10146243125 | -0.1014613584 | -0.10146123915 |
| α_3 | -10.00566005706 | -10.00646209716 | -10.0065460205 | -10.0065546035 |
| α_4 | 1.002233743668 | 1.002198219299 | 1.00219523907 | 1.002194881439 |
| α_5 | 12.82191181182 | 12.82294845581 | 12.8230447769 | 12.82305812835 |
| α_6 | 73.40586090087 | 73.40718841552 | 73.4072723389 | 73.40727996826 |

从上面的两个表上就可以清楚地看出，跟最小二乘模型相比，待定系数随输入误差的量级增大改变的要比最小二乘模型要小一些，变分同化方法明显更稳定。特别是当输入误差的量级大到 10^0 的时候，最小二乘模型的待定系数就已经完全改变了，而变分同化待定系数改变不是很大，才只是 10^{-1} 量级。充分说明了变分同化模型再对于输入值含有误差问题上更稳定，更有优势！

十、问题 II -4—中央差商的最小二乘法模型

此问题中不仅观测资料有误差，连时间变量也含有误差，试利用数据 DATA4.TXT，建立数学模型，确定参数 α_k ($1 \leq k \leq 6$) 在某种意义下的最优解。

10.1 分析和模型修改

如果观测资料的时间变量 t_j 也含有误差，如 Data4.txt 中的数据，此时时间序列不是等间距的，则原线性方程组的差分格式中的导数差分不能采用问题三中的高精度的差分格式，即矢量方程 $A\bar{\alpha} = b$ 的右端向量 \bar{b} 要改变。经实践证明，

如果仍然采用一阶精度的后差替换,再用与问题三相同的方法计算出来的参数来模拟,效果不是很好。为了提高精度,在此改用中央差公式替换导数项。则有:

$$b = \begin{bmatrix} \frac{x(t_{i+1}) - x(t_{i-1}))}{t_{i+1} - t_{i-1}} \\ \frac{y(t_{i+1}) - y(t_{i-1}))}{t_{i+1} - t_{i-1}} \\ x(1) \\ y(1) \end{bmatrix}_{(2n+2) \times 1}$$

其中, $i = 2, 4, \dots, m-1; n = m-2$;

因为此时 data4.txt 的时间间隔短,数据密度大,中央差可以很好地提高计算精度。通过程序运行的结果和模拟的情况来看,仿真结果确实得到提高。

模拟仿真时仍然采用龙格库塔方法。

10.2 结果讨论

对于 DATA4 数据,由上面的分析,再用最小二乘法编写程序四求出参数 α_i , 为

$$\alpha_1 = 1.6497825740, \alpha_2 = -0.0901167975, \alpha_3 = -12.2486331003$$

$$\alpha_4 = 1.1221869484, \alpha_5 = 13.2120617404, \alpha_6 = 70.8366426217$$

将参数 α_i 带入原方程组,采用龙格库塔方法模拟出的结果来看,模拟值和观测数据吻合的非常好,数据的相关系数达到 99.484%和 99.687%。参考图 20, 图 21。

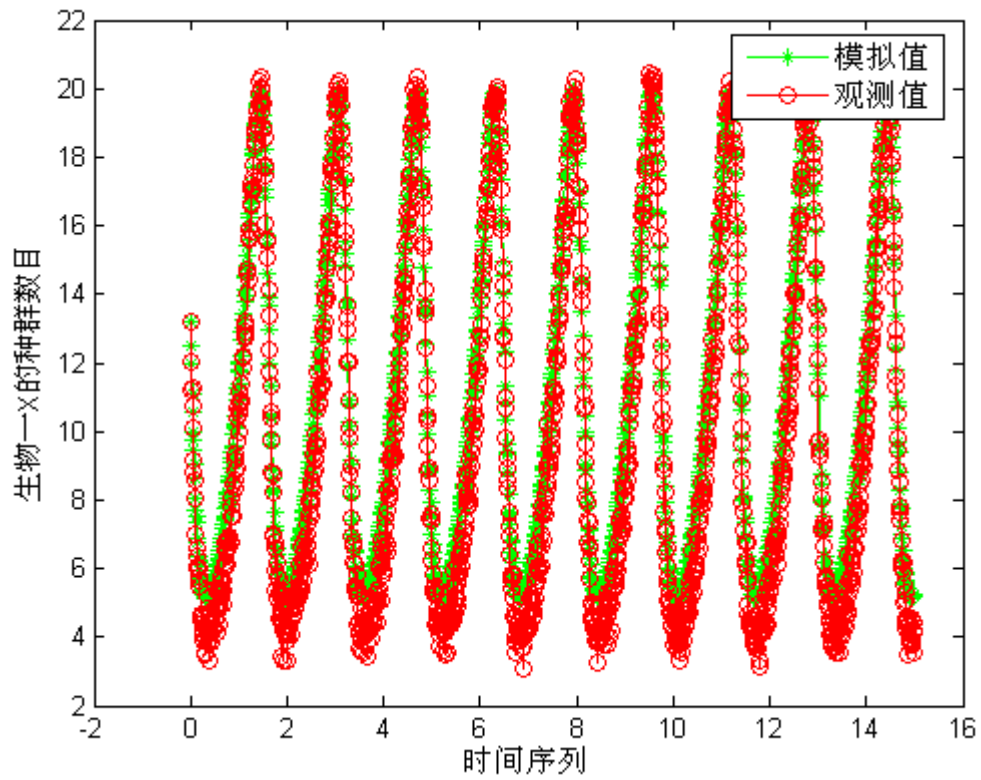


图 20 龙格库塔方法对种群 X 的模拟值和观测值的比较图

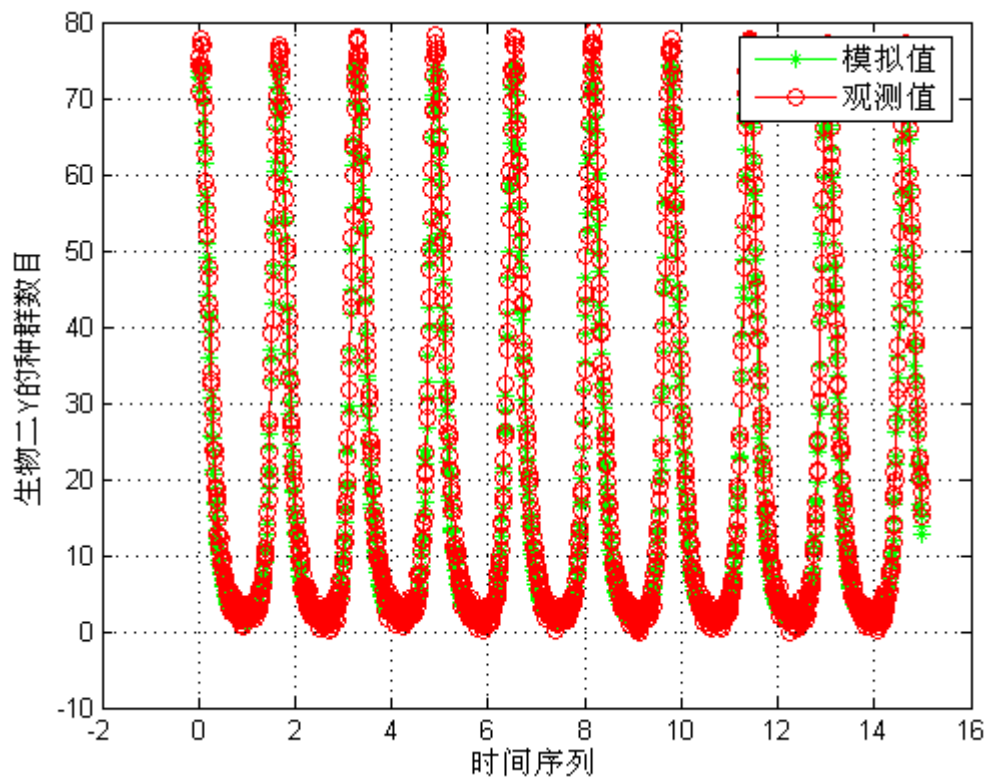


图 21 龙格库塔方法对种群 Y 的模拟值和观测值的比较图

通过对 DATA4 的仿真，我们发现用中央差商的最小二乘法也能够仿真真实

值相当的好。但是由于题目中说了既有观测资料的误差还有时间的误差，所以我们更要对其误差敏感度进行检验。

10.3 对模型误差敏感度的检验：

与前面的 8.3.3 一样，我们也要对此问中的最小二乘模型进行误差敏感度的检验。但是此问又与上问不同，这时除了观测资料有误差外，连时间变量也含有误差，这就说明对模型敏感度的要求更高，需要模型更加的稳定。

对于此题中的最小二乘模型，对于 DATA4 数据，我们也还将观测资料和时间变量的输入的随机误差量级分别从 10^0 到 10^{-6} 进行讨论，可以看出其导致的模型求解的待定系数变化如下表所示：

表 11 DATA4 输入误差对最小二乘模型待定系数的关系表

| <div>误差量级</div> <div>系数</div> | 10^0 | 10^{-1} | 10^{-2} | 10^{-3} |
|-------------------------------|---------------|---------------|---------------|---------------|
| α_1 | 0.1624077741 | 0.0508724178 | 2.0544677112 | 2.3582404146 |
| α_2 | 0.0005175474 | 0.0030743333 | -0.0452105832 | -0.0832197710 |
| α_3 | 0.0074801606 | 0.3028667305 | -1.9389988538 | -4.0727520174 |
| α_4 | 0.0033648823 | -0.0369842222 | 0.1380542123 | 0.1070515691 |
| α_5 | -4.4650067337 | 11.7734855448 | 13.0133618620 | 13.2065655746 |
| α_6 | 61.1784896666 | 69.2219560092 | 70.7094834422 | 70.8274408283 |

| 误差量级 系数 | 10^{-4} | 10^{-5} | 10^{-6} | 无误差 |
|------------|---------------|----------------|----------------|----------------|
| α_1 | 1.6415072989 | 1.6522005137 | 1.6492999845 | 1.6497825740 |
| α_2 | -0.0856078779 | -0.0911924751 | -0.0899122845 | -0.0901167975 |
| α_3 | -8.8745751609 | -12.8712449615 | -12.1167133337 | -12.2486331003 |
| α_4 | 0.8940819240 | 1.1659759803 | 1.1129487696 | 1.1221869484 |
| α_5 | 13.2106818717 | 13.2119410649 | 13.2120491557 | 13.2120617404 |
| α_6 | 70.8351935209 | 70.8365132801 | 70.8366273042 | 70.8366426217 |

虽然时间变量也有了误差，但是从表上可以看出，在输入误差的量级小于 10^{-2} 的时候，我们模型的待定系数仍然是变化不大的，可见模型的稳定性还是很好的。如图 22，图 23 所示是当误差量级为 10^{-4} 时的仿真模拟图。从中可以看出，模拟值和观测值吻合的非常好。

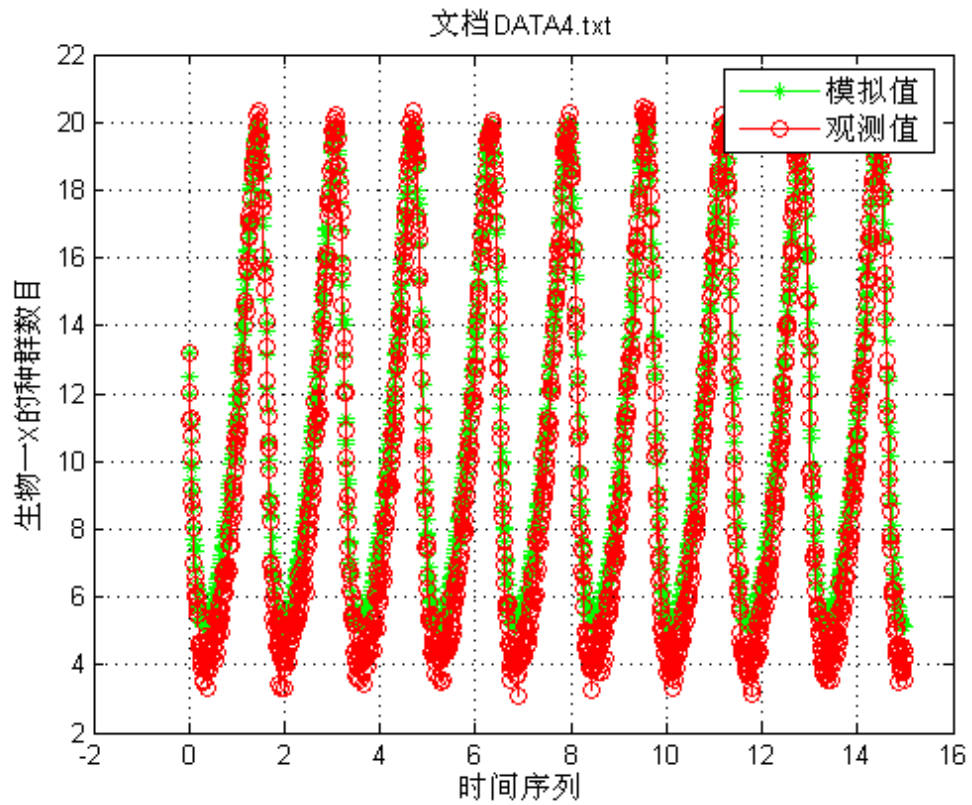


图 22 当误差量级为 10^{-4} 时，龙格库塔方法对种群 X 的模拟图

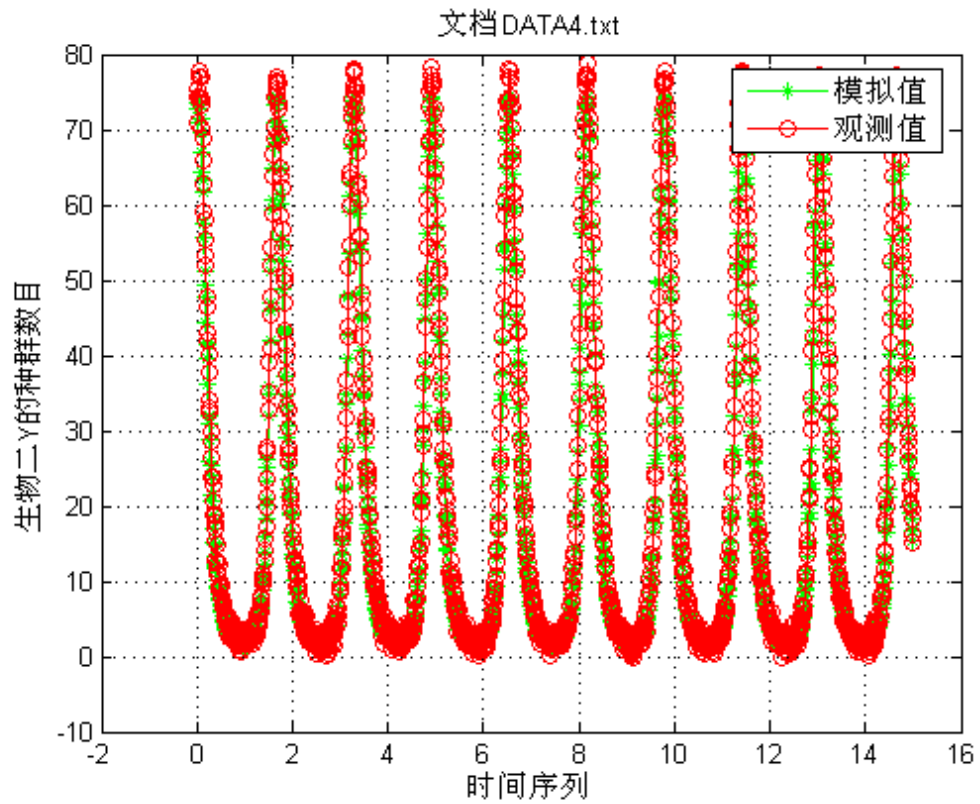


图 23 当误差量级为 10^{-4} 时，龙格库塔方法对种群 Y 的模拟图

但是当输入误差的量级大于 10^{-2} 的时候，时间变量误差就体现出来其影响了，这时模型待定系数的改变量明显要大于问题三，改变很大。这说明在模型的稳定性上还需要改进一点，让其更稳定，我们仍然用变分伴随同化来改进。

十一、问题 II -4—变分伴随同化模型

11.1 问题的分析

当观测时间也存在误差时，我们发现 data4 种的数据各个数据之间时间间隔并不是等距的。此时需要对差分和积分格式作相应的修改。

11.2 问题的求解：

根据上面的分析，对第三问中变分伴随同化方法的程序中时间差分，积分格式作相应的修改，然后对 data4 的数据进行分析，得到结果如下。

表 12 给出了采用变分伴随同化方法对参数进行优化前后的参数，其中优化前的参数为最小二乘法模型计算得到的参数值。从表 13 中可以看出，优化后相关系数有所提高，平均绝对误差也有所减小。从图中可以看出，优化前后的仿真值都能够很好地反映种群数目的变化趋势。从图 26 上则可以看出，优化后仿真结果的误差比较小。对 A 种群而言，误差在 $[-2.0, 2.0]$ 范围内，B 种群，误差在 $[-6.0, 2.0]$ 范围内。从相空间平面图上也可以看出，仿真结果与观测值比较吻合。

表 12 优化前后参数对比

| 参数 | 优化前 | 优化后 |
|------------|----------------|-----------------|
| α_1 | 1.6497825740 | 2.079206875230 |
| α_2 | -0.0901167975 | -0.096655866585 |
| α_3 | -12.2486331003 | -9.368555302945 |
| α_4 | 1.1221869484 | 0.958584040818 |
| α_5 | 13.2120617404 | 13.212061882019 |
| α_6 | 70.8366426217 | 70.836639404297 |

表 13 优化前后仿真值与观测值的相关系数和平均绝对误差

| 相关系数 | 优化前 | 优化后 |
|------|----------------|----------------|
| A种群 | 0.996631995780 | 0.996685204494 |
| B种群 | 0.997822624574 | 0.998049234450 |

| 平均绝对误差 | 优化前 | 优化后 |
|--------|----------------|----------------|
| A种群 | 0.405199100989 | 0.404988184656 |
| B种群 | 1.721126213560 | 1.717043180639 |

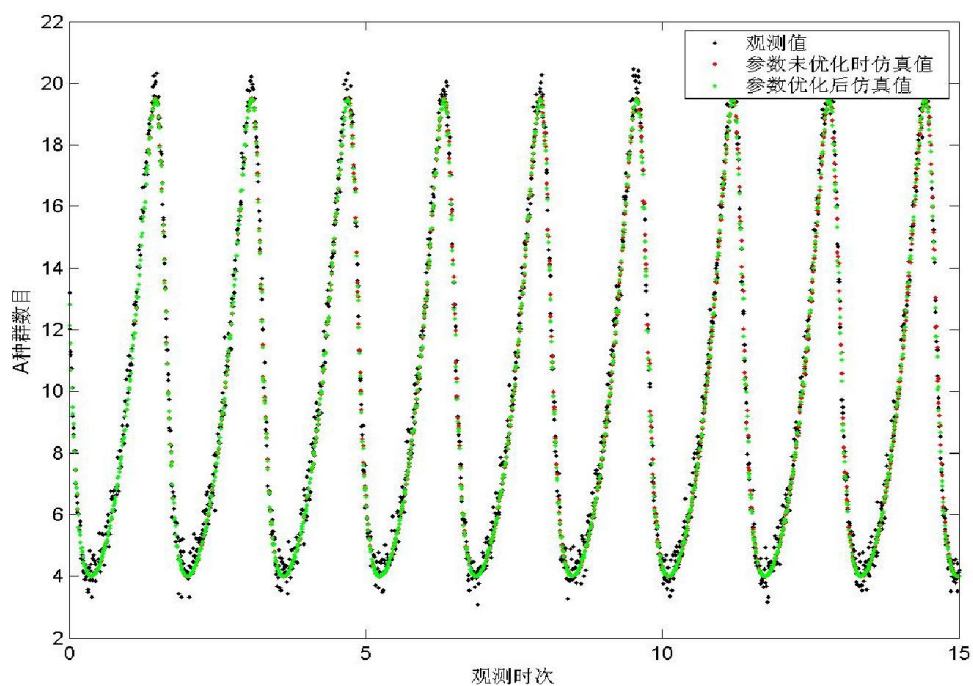


图 24 A 种群数目的观测值、未优化参数计算的仿真值、优化后参数计算的仿真值比较图

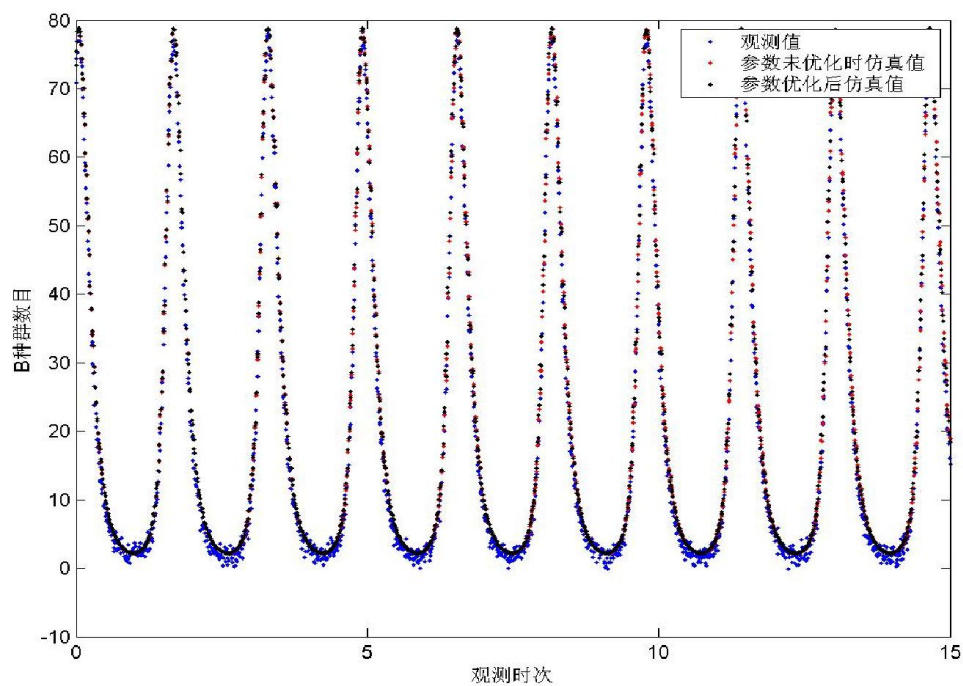


图 25 B 种群数目的观测值、未优化参数计算的仿真值、优化后参数计算的仿真值比较图

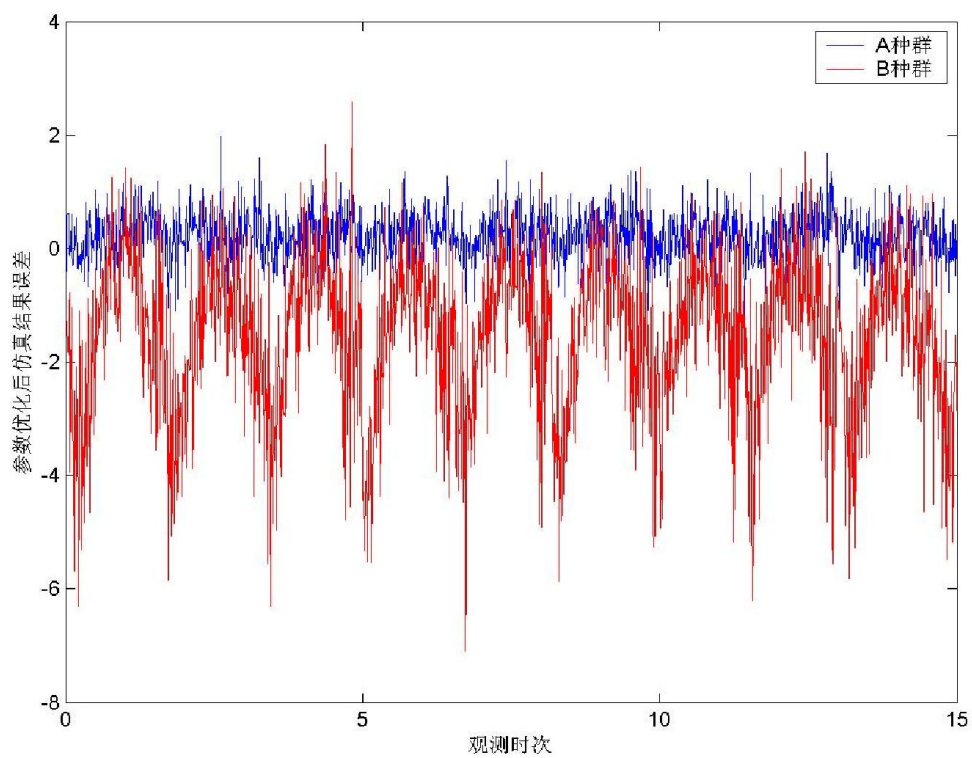


图 26 参数优化后仿真结果误差

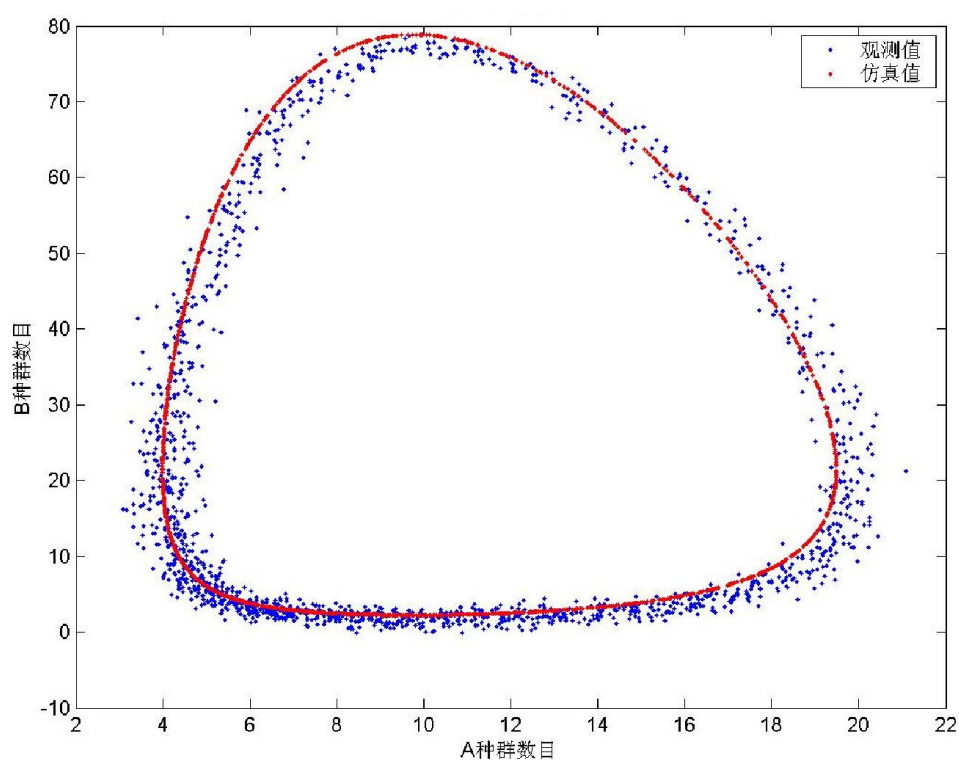


图 27 相平面图

11.3 对模型误差敏感度的检验：

与前面的 10.3 一样，我们也要对此问中的变分同化模型进行误差敏感度的检验，同样对 DATA4 我们这里也既要考虑观测资料有误差外，也要考虑时间变量也含有误差。也还将观测资料和时间变量的输入的随机误差量级分别从 10^0 到 10^{-6} 进行讨论，可以看出其导致的模型求解的待定系数变化如下表所示：

表 14 DATA4 输入误差对变分同化模型待定系数的关系表

| 误差量级 系数 | 10^0 | 10^{-1} | 10^{-2} | 10^{-3} |
|------------|-----------------|-----------------|-----------------|-----------------|
| α_1 | 2.529336160377 | 2.074853641746 | 2.083425004937 | 2.079117145436 |
| α_2 | -0.365517353010 | -0.144805502260 | -0.094273794126 | -0.096262217054 |
| α_3 | -9.261712719403 | -9.336414586515 | -9.371792641500 | -9.368997411640 |
| α_4 | 0.944566509527 | 0.953054377253 | 0.9576411029486 | 0.958436908950 |
| α_5 | 13.603360848167 | 13.223605116829 | 13.216416579010 | 13.212375048516 |
| α_6 | 71.098736237324 | 70.865833108039 | 70.840808448696 | 70.836149265597 |

| 误差量级 系数 | 10^{-4} | 10^{-5} | 10^{-6} | 无误差 |
|------------|-----------------|-----------------|-----------------|-----------------|
| α_1 | 2.079170764318 | 2.079202027969 | 2.079207221451 | 2.079206875230 |
| α_2 | -0.096685590063 | -0.096653398728 | -0.096655841432 | -0.096655866585 |
| α_3 | -9.368585430770 | -9.368555851980 | -9.368555600297 | -9.368555302945 |
| α_4 | 0.958594420065 | 0.958588358963 | 0.958584212955 | 0.958584040818 |
| α_5 | 13.212039100811 | 13.212061541962 | 13.212062220137 | 13.212061882019 |

| | | | | |
|------------|-----------------|-----------------|-----------------|-----------------|
| α_6 | 70.836609285723 | 70.836638590791 | 70.836638923936 | 70.836639404297 |
|------------|-----------------|-----------------|-----------------|-----------------|

虽然时间变量也有了误差，但是从表上可以看出，即使在输入误差的量级大于 10^{-2} 的时候，我们模型的待定系数仍然是变化不大的，可见模型的稳定性还是相当好的。特别是和 8.4 的最小二乘的量级大于 10^{-2} 部分相比，我们系统稳定性的优势就很明显能看出来了，这样即使对于观测资料和时间变量都有误差的情况下，我们这种方法仍然是很稳定的。

十二、模型的推广

此生态系统还是比较简单的系统，在实际问题中，我们还是要考虑更加周全一点，所以必须对模型再作进一步推广。

12.1 背景知识:

由原来描述捕食者和被捕食者关系的微分方程组，即：

$$\begin{cases} \frac{dx}{dt} = x[\alpha_1 + \alpha_2 y] \\ \frac{dy}{dt} = y[\alpha_3 + \alpha_4 x] \end{cases}$$

进行变换可以得到：

$$\begin{cases} x = -\frac{\alpha_3}{\alpha_4} + \frac{1}{y} \frac{dy}{dt} \\ y = -\frac{\alpha_1}{\alpha_2} - \frac{1}{x} \frac{dx}{dt} \end{cases}$$

对微分方程组每一个式子两边从 0 到 T 积分得到：

$$\begin{cases} x_{mean} = \frac{1}{T} \int_0^T x(t) dt = -\frac{\alpha_3}{\alpha_4} + \int_0^T \left(\frac{1}{y} \frac{dy}{dt} \right) dt \\ y_{mean} = \frac{1}{T} \int_0^T y(t) dt = -\frac{\alpha_1}{\alpha_2} - \int_0^T \left(\frac{1}{x} \frac{dx}{dt} \right) dt \end{cases}$$

最终得到两类生物数量的平均值为：

$$\begin{cases} x_{mean} = -\frac{\alpha_3}{\alpha_4} + \frac{1}{T}(\ln y(T) - \ln y(0)) \\ y_{mean} = -\frac{\alpha_1}{\alpha_2} - \frac{1}{T}(\ln x(T) - \ln x(0)) \end{cases}$$

但是由于 T 是 $x(t)$ 的周期，从而有：

$$\ln x(T) - \ln x(0) = 0$$

类似可以得到：

$$\ln y(T) - \ln y(0) = 0$$

那么，两类生物数量的平均值分别是：

$$\begin{cases} x_{mean} = -\frac{\alpha_3}{\alpha_4} \\ y_{mean} = -\frac{\alpha_1}{\alpha_2} \end{cases}$$

12.2 考虑捕获量这个外界因素的影响：

实际中，除了 A 生物和 B 生物互相之间的制约关系外，比如还要考虑在原有模型中添加人为捕杀的因素，不妨设捕获率跟生物数量成正比，则此时 A 生物和 B 生物的增长的微分方程组变为：

$$\begin{cases} \frac{dx}{dt} = x(t)[\alpha_1 + \alpha_2 y(t)] - hx(t) = x(t)[\alpha_1 - h + \alpha_2 y] \\ \frac{dy}{dt} = y(t)[\alpha_3 + \alpha_4 x(t)] - hy(t) = y(t)[\alpha_3 - h + \alpha_4 x(t)] \end{cases}$$

类似于前面的背景知识分析可以得到，有捕获时，两类生物的数量仍然呈周期性变化，而且在一个周期内两类生物的数量平均值分别为：

$$\begin{cases} x_{mean} = -\frac{\alpha_3}{\alpha_4} - \frac{h}{\alpha_4} \\ y_{mean} = -\frac{\alpha_1}{\alpha_2} + \frac{h}{\alpha_2} \end{cases}$$

这表明：在相同的资源下，减少捕获量即 h 的减小，被捕食者 B 的平均数量也相应减小，而捕食者 A 的数量却因此会增加。因此，减少捕获量对捕食者 A 的种群数量的增加有利，而对被捕食者 B 的种群增长不利。

再推广一下，如果将这种模型的研究结果应用到农作物的害虫和它的天敌系统，可能得到意想不到的结论。

将 x 设为农作物的害虫，而将 y 设为此种害虫的天敌。若使用某种杀虫剂来灭虫，但这种杀虫剂不仅能杀死害虫，也能杀死其天敌。因此，杀虫剂的使用相当于上面模型中的捕获量。这样，由捕食与被捕食系统的规律可知，用的杀虫剂越多即捕获量越大，对天敌的数量增长不利，而对害虫有利。因此，在用杀虫剂灭虫时，应该适当用量，而并非多多益善。

对于更多个生物群体所形成的群落，各种群间的相互作用的情况要比两种群的情况复杂。虽然每两个种群之间就是相互竞争，相互共存和捕食与被捕食这些基本关系，但由于多种种群的两两关系的不同，会产生种类繁多的关系，但建立模型的思路基本上与两种群情况相同：即既要考虑种群自身的增长规律，又要考虑种群的相互作用。

12.3 考虑更为广泛的捕食系统：

在实际中，由于还有很多更加复杂的情况。根据参考文献[3]中所说，还有更为广泛的捕食系统，所以我们有必要讨论一下更为广泛的捕食系统：

$$\begin{aligned}\frac{dx}{dt} &= b_1x + b_2x^2 + b_3x^3 + b_4xy \\ \frac{dy}{dt} &= b_5y + b_6xy - b_7y^2\end{aligned}\tag{12}$$

然后还是用问题四中用到的最小二乘法来求最优的系数，（用的数据是

DATA3),具体过程这里就不细说了,编写程序八,可以求出此时最优的系数是:

$$b_1 = 2.0906, b_2 = -0.0031338, b_3 = 5.2108 \times 10^{-5}, b_4 = -0.10348,$$

$$b_5 = -10.232, b_6 = 1.022, b_7 = 0.00045453$$

把最优系数带回方程组(12),然后跟问题四一样方法仿真出 $t = 0.1$ 到 $t = 15$ 的

(\bar{X}, \bar{Y}) ,跟观测值进行比较,结果如图:

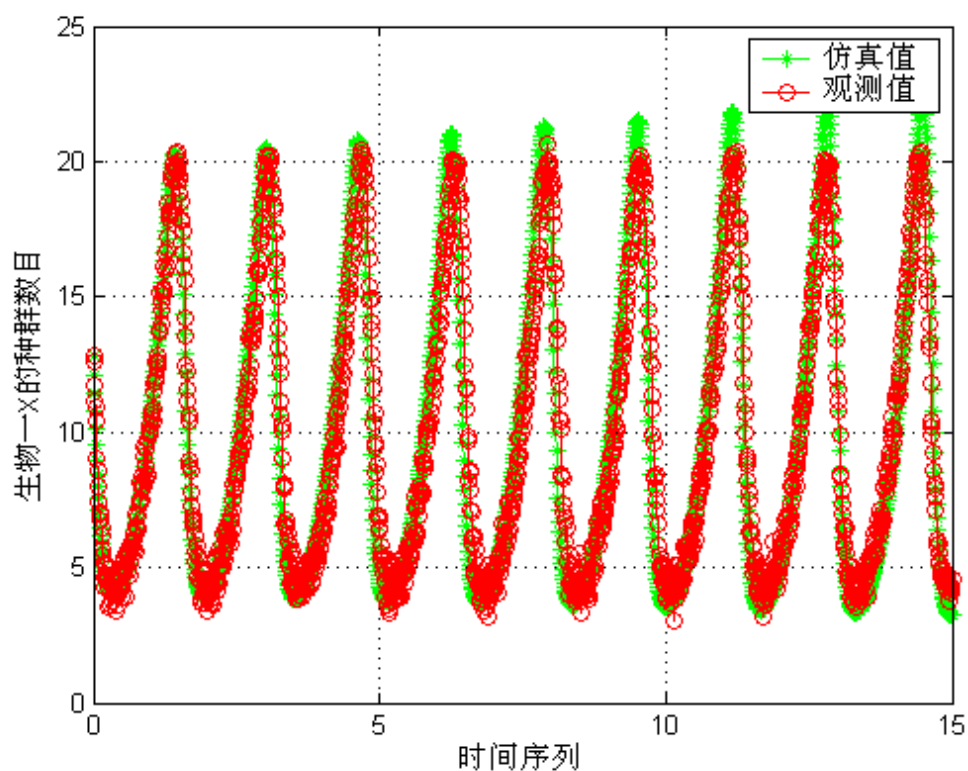


图 28 A 种群的仿真值与真实值的比较图

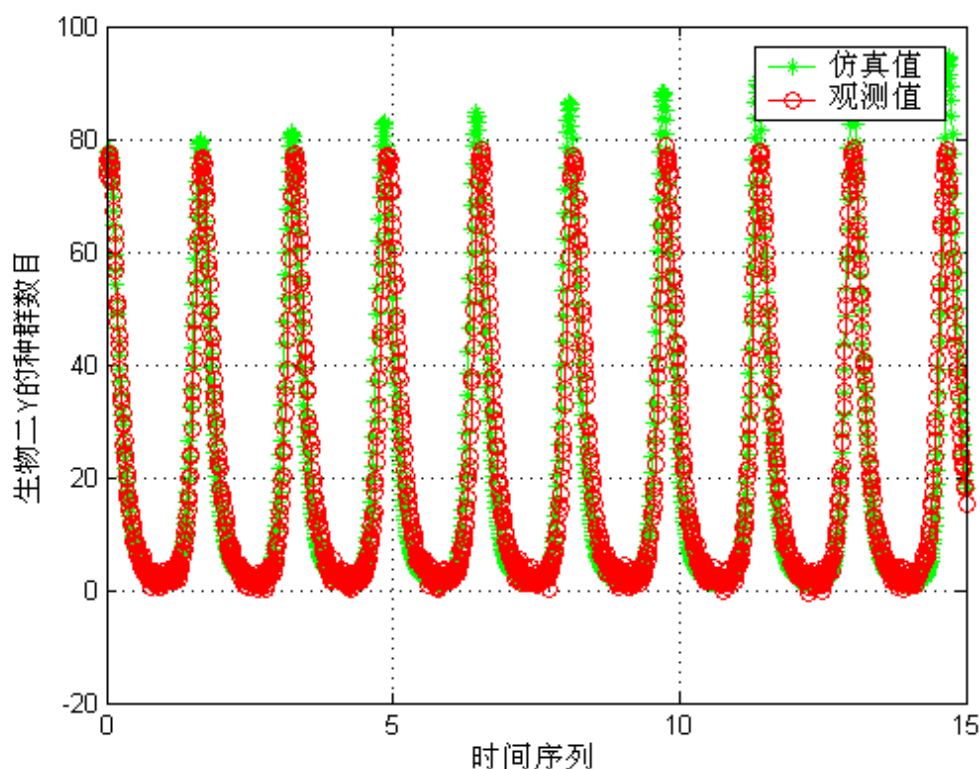


图 29 B 种群的仿真值与真实值的比较图

可以看出效果还是比较好的,两种种群的仿真值相和观测值的相关系数分别是 0.96002 和 0.93953。发现两点:

约束条件一多,待定参数一多,仿真值的效果就稍微差了一点.这是与实际相符合的.一旦考虑的约束条件多,情况变复杂,那么就越来越不容易得到较好的仿真值。

看求出得参数的量级，可以看出 b_2 ， b_3 ， b_7 的量级相比较而言比较小，所以近似可以略去，如果略去，剩下的方程组就是：

$$\begin{aligned}\frac{dx}{dt} &= b_1x + b_4xy \\ \frac{dy}{dt} &= b_5y + b_6xy\end{aligned}$$

可以看出这样与题目所给的方程组一样的，可以题目所给的这个方程组还是对实际问题很合理的简化。

十三、模型的优缺点分析

优点：

1. 通过实际检验，发现最小二乘模型在本题中是比较稳定的；
2. 通过比较试验，发现最小二乘法与变分同化伴随方法的结论及观测资料比较接近；
3. 对于非线性或者变系数的情况下，可以再变分同化的基础上加上正则化方法改进。
4. 模型的适用范围广。

缺点：

1. 利用数值差分格式构造二元线性方程组的过程中存在误差，若给出的原始资料时间步长较大，模型误差就大，会导致由此求出的参数不真实；
2. 模型的精度依赖于数值差分的精度和初始资料的精度较明显；
3. 利用解析解的情况下，积分常数 K 的确定存在误差，对结论有一定影响；

十四、关于本问题的几点讨论

1. 参数反演往往是不适定和非线性，，尤其是不适定性使得通常的数值方法失败。对于一个系统，通常有三部分组成：输入信号 x ，模型 K ，输出信号 g ，于是，三者的关系用如下的数学模型来表示

$$Kx = g$$

已知 g 求 x 一般称为反问题，这类问题往往是不适定的，对于不适定的问题，我们把精力放在解的稳定性上，设法构造稳定性近似解来逼近“真解”。数据的观测误差容易造成系数变化大，故在最小二乘法中，要使 $\|Ax - b\|^2 = \min!$ ，若行列式 $\|A^T A\| \approx 0$ ，模型的性能尤为重要，此时宜采用正则化方法处理。

可以引入 Tikhonov 函数

$$J[\alpha] = \|A_h \alpha - b^\delta\|^2 + \alpha \|x\|^2 = \min!$$

其中 A_h, b^δ 为带误差的近似值（因为 A 中包含了观测资料的误差）。适

定选择 δ 和 h 的条件

$$\begin{cases} \|A_h - A_{\text{真}}\| \leq h \\ \|b^\delta - b_{\text{真}}\| \leq \delta \end{cases}$$

使得当 $\delta \rightarrow 0, h \rightarrow 0$ 时, $\alpha^\delta \rightarrow \alpha^*$ (真值), 则此时的模型是稳定的。

此例中的观测误差计算得到

$$\|A^T A\| \neq 0$$

2. 最小二乘法具有一定的精度。实际上, 如果参数为随时间而变化的, 即 $\alpha_i(t)$, 此时最小二乘方法不能体现出优越性, 而变分同化方法的效果将更加明显。在模型的拓展中我们将介绍如何改进前面的变分伴随同化方法, 解决参数随时间变化等问题。

3. 观测数据不能太多, 如果观测数据给的太多, 处理数据费时, 也不经济。实际上当观测资料数据给出到一定数量对本题特别重要, 亦即说观测资料应该给多少这对参数反演更为重要。或者说, 在给出的众多数据中, 提取多少的数据量居足以在较高的精度内反演出参数。

十五、模型的扩展

参数反演是一个很复杂的问题。题目中建立的是比较简单的捕食者—被捕食者生态模型, 当面对更复杂的生态模型时, 如何确定其中参数呢? 我们使用的方法能不能用来求解这些问题呢? 这是一个比较有意义的问题。本文主要采用了两种方法: 最小二乘法和变分同化方法。考虑更富杂的生态模型时, 变分同化的方法比最小二乘法更有优越性。在本文提供的生态模型的基础上, 我们进行一些拓展, 并给出如何采用变分同化方法进行求解的思路。由于时间有限, 我们没有进行数值模拟。

1. 当参数 α_i 随时间变化时, 此时建立的泛函形式仍然不变, 但是泛函梯度 $\nabla_{\alpha_i} J$ 的计算公式发生了改变, 如下所示:

$$\begin{cases} \nabla_{\alpha_1} J = xP \\ \nabla_{\alpha_2} J = xQ \\ \nabla_{\alpha_3} J = xyP \\ \nabla_{\alpha_4} J = xyQ \\ \nabla_{\alpha_5} J = P(t_0) \\ \nabla_{\alpha_6} J = Q(t_0) \end{cases}$$

2. 当考虑其他的影响因素，方程组右边加入控制项 $f_1(t)$, $f_2(t)$ ，即：

$$\begin{cases} \dot{x} = x(\alpha_1 + \alpha_2 y) + f_1(t) \\ \dot{y} = y(\alpha_1 + \alpha_2 x) + f_2(t) \end{cases}$$

此时建立的目标泛函同样形式不变，但要给出 $\nabla_{f_1} J$ 和 $\nabla_{f_2} J$ 。推导得到：

$$\begin{cases} \nabla_{f_1} J = P(t) \\ \nabla_{f_2} J = Q(t) \end{cases}$$

对于上述两种情况，对于给定的初值，都可以利用文中变分同化方法的流程进行寻求最优参数。

参考文献：

- [1] 大气科学中的数学物理问题，黄思训、伍容生编著，气象出版社，2001
- [2] 数学试验，重庆大学数学系主编，科学出版社，2001
- [3] 一类捕食者—被捕食者种群模型的定性分析，王元明、黄汛成，湛江海洋大学学报，2005，vol25，no4
- [4] Matlab 6 工程计算及应用，何仁斌编著，重庆大学出版社，2001
- [5] MATLAB 遗传算法工具箱及应用，雷英杰、张善文等，西安电子科技大学出版，2005
- [6] 遗传算法与工程优化，玄光男、程润伟等著译，清华大学出版社
- [7] 概率论与数理统计，四川大学数学学院组编，刘晓石、陈鸿建、何腊梅编著，科学出版，200
- [8] 数学建模案例精选，朱道元等编著，科学出版社，2003
- [9] 数学手册，潘志、贺祖琪主编，中国矿业大学出版，1995